

# **A SMALL LOW-POWER WIRELESS INTEGRATED MICROSYSTEM**

by

Bennion Redd

A dissertation submitted to the faculty of  
The University of Utah  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

The University of Utah

August 2015

Copyright © Bennion Redd 2015

Some Rights Reserved (See Appendix)

# The University of Utah Graduate School

## STATEMENT OF DISSERTATION APPROVAL

The dissertation of Bennion Redd  
has been approved by the following supervisory committee members:

<u>Richard B. Brown</u>	, Chair	<u>March 16, 2015</u> Date Approved
<u>Erik Brunvand</u>	, Member	<u>March 16, 2015</u> Date Approved
<u>Kent Smith</u>	, Member	<u>March 16, 2015</u> Date Approved
<u>Kenneth Stevens</u>	, Member	<u>March 16, 2015</u> Date Approved
<u>Ling Zang</u>	, Member	<u>March 16, 2015</u> Date Approved

and by Gianluca Lazzi, Chair/Dean of  
the Department/College/School of Electrical and Computer Engineering

and by David B. Kieda, Dean of The Graduate School.

## ABSTRACT

Advancements in process technology and circuit techniques have enabled the creation of small chemical microsystems for use in a wide variety of biomedical and sensing applications. For applications requiring a small microsystem, many components can be integrated onto a single chip. This dissertation presents many low-power circuits, digital and analog, integrated onto a single chip called the Utah Microcontroller. To guide the design decisions for each of these components, two specific microsystems have been selected as target applications: a Smart Intravaginal Ring (S-IVR) and an NO releasing catheter. Both of these applications share the challenging requirements of integrating a large variety of low-power mixed-signal circuitry onto a single chip. These applications represent the requirements of a broad variety of small low-power sensing systems.

In the course of the development of the Utah Microcontroller, several unique and significant contributions were made. A central component of the Utah Microcontroller is the WIMS Microprocessor, which incorporates a low-power feature called a scratchpad memory. For the first time, an analysis of scaling trends projected that scratchpad memories will continue to save power for the foreseeable future. This conclusion was bolstered by measured data from a fabricated microcontroller. In a 32 nm version of the WIMS Microprocessor, the scratchpad memory is projected to save  $\sim 10\text{--}30\%$  of memory access energy depending upon the characteristics of the embedded program. Close examination of application requirements informed the design of an analog-to-digital converter, and a unique single-opamp buffered charge scaling DAC was developed to minimize power consumption. The opamp was designed to simultaneously meet the varied demands of many chip components to maximize circuit reuse.

Each of these components are functional, have been integrated, fabricated, and tested. This dissertation successfully demonstrates that the needs of emerging small low-power microsystems can be met in advanced process nodes with the incorporation of low-power circuit techniques and design choices driven by application requirements.

# CONTENTS

<b>ABSTRACT</b> .....	<b>iii</b>
<b>LIST OF TABLES</b> .....	<b>vii</b>
<b>LIST OF FIGURES</b> .....	<b>viii</b>
<b>ACKNOWLEDGMENTS</b> .....	<b>xi</b>
<b>CHAPTERS</b>	
<b>1. APPLICATIONS OF MIXED-SIGNAL MICROSYSTEMS</b> .....	<b>1</b>
1.1 Introduction .....	1
1.2 Unique Contributions .....	2
1.3 Related Microsystems .....	3
1.4 Smart Intravaginal Ring .....	5
1.4.1 AIDS and the Promise of Microbicides .....	5
1.4.2 User Adherence for Microbicide Trials .....	6
1.4.3 User Adherence Studies .....	6
1.4.4 Fertility Assistance .....	7
1.4.5 Physical Description of S-IVR .....	7
1.5 NO-releasing Catheter .....	8
1.6 Common Microsystem Requirements .....	11
<b>2. UTAH MICROCONTROLLER</b> .....	<b>14</b>
2.1 Need for a Custom Microcontroller .....	14
2.2 Background on the WIMS Microprocessor .....	16
2.3 Architecture for Increased Address Space in an Ultra-Low-Power Microprocessor .....	18
2.3.1 Introduction .....	18
2.3.2 Measurements .....	20
2.3.3 Minimizing Code Size .....	23
2.3.4 Alternatives .....	24
2.3.5 Conclusion .....	25
<b>3. SCRATCHPAD MEMORY</b> .....	<b>26</b>
3.1 Introduction .....	26
3.2 Background .....	27
3.3 Motivation .....	28
3.4 Conceptual Framework .....	28
3.4.1 Models .....	28
3.4.2 Relative Importance of Active and Leakage Energy .....	29

3.4.3	Access Energy and Memory Size . . . . .	30
3.4.4	Leakage Energy and Process Scaling . . . . .	31
3.5	Scaling Trends . . . . .	32
3.5.1	Process Innovations . . . . .	33
3.5.2	SRAM Cell Design . . . . .	35
3.6	WIMS Microcontroller . . . . .	35
3.6.1	Measured Data . . . . .	36
3.6.2	32 nm Projections . . . . .	38
3.6.3	WIMS Scratchpad Memory Approach . . . . .	40
3.6.4	Static Benchmarks . . . . .	42
3.6.5	Dynamic Benchmarks . . . . .	43
3.7	Conclusions . . . . .	45
<b>4.</b>	<b>A LOW-POWER 3-STAGE 1.8 V OPERATIONAL AMPLIFIER . . . .</b>	<b>46</b>
4.1	Introduction . . . . .	46
4.2	Desired Specifications . . . . .	46
4.3	Opamp Architecture . . . . .	47
4.3.1	Limitations of Classic 2-stage Opamp . . . . .	47
4.3.2	A Suitable Input Stage . . . . .	49
4.3.3	A Suitable Output Stage . . . . .	49
4.3.4	Opamp Schematic and Operation . . . . .	50
4.3.5	Noise Analysis . . . . .	53
4.4	Simulated and Measured Results . . . . .	55
4.5	Conclusion . . . . .	59
<b>5.</b>	<b>A LOW-POWER HIGH-PRECISION SIGMA-DELTA ADC . . . . .</b>	<b>60</b>
5.1	Introduction . . . . .	60
5.2	Desired Specifications . . . . .	61
5.3	ADC Design . . . . .	64
5.3.1	Architecture Selection . . . . .	64
5.3.2	Low-power Sigma-delta Literature Review . . . . .	66
5.3.3	Sigma-delta Modulator Design . . . . .	67
5.3.4	Decimation Filter Design . . . . .	73
5.4	Results . . . . .	74
5.4.1	Simulation Parameter Selection . . . . .	74
5.4.2	Simulated Performance of the Modulator . . . . .	75
5.4.3	Simulated Performance of the Decimation Filter . . . . .	75
5.4.4	Measured Results . . . . .	76
5.4.5	Result Summary . . . . .	80
5.5	Conclusion . . . . .	80
<b>6.</b>	<b>A LOW-POWER SINGLE-OPAMP CHARGE SCALING 10-BIT DAC</b>	<b>82</b>
6.1	Introduction . . . . .	82
6.2	Desired Specifications . . . . .	82
6.3	DAC Background . . . . .	83
6.3.1	Common DAC Architectures . . . . .	83
6.3.2	Low-power DAC Literature Review . . . . .	87
6.4	DAC Design . . . . .	88
6.4.1	Capacitor Sizing . . . . .	89
6.5	Simulated Performance . . . . .	90

6.5.1 Simulation Result Summary . . . . .	92
6.6 Measured Performance . . . . .	93
6.7 Conclusion . . . . .	97
<b>7. CONCLUSION . . . . .</b>	<b>98</b>
7.1 Motivation . . . . .	98
7.2 Dissertation Accomplishments . . . . .	98
7.3 Future Work . . . . .	103
7.4 Conclusion . . . . .	104
<b>APPENDIX: CREATIVE COMMONS ATTRIBUTION 4.0 INTER-</b>	
<b>NATIONAL PUBLIC LICENSE . . . . .</b>	<b>106</b>
<b>REFERENCES . . . . .</b>	<b>113</b>

## LIST OF TABLES

2.1 Commercial microcontrollers and application requirements. . . . .	16
4.1 Opamp requirements. . . . .	47
4.2 Simulated and measured opamp results. . . . .	56
5.1 Sensor data and time resolution requirements. . . . .	61
5.2 ADC requirements. . . . .	64
5.3 Characteristics of several example low-power ADCs. . . . .	66
5.4 ADC simulation parameters. . . . .	75
5.5 ADC simulation results. . . . .	77
5.6 ADC measured results. . . . .	80
6.1 DAC specifications. . . . .	83
6.2 DAC switch timing. . . . .	89
6.3 DAC simulated and measured results. . . . .	93



## LIST OF FIGURES

1.1 A millimeter-size microsystem. . . . .	4
1.2 A microsystem with a pH meter, transmitter, and battery. . . . .	4
1.3 Depiction of finished S-IVR. . . . .	8
1.4 Commercial sensor chip provided by e-SENS. . . . .	9
1.5 Clotting prevention of NO-releasing catheters. . . . .	10
1.6 Diagram of NO-releasing catheter. . . . .	11
1.7 NO-releasing catheter microsystem. . . . .	12
2.1 Layout of the third-generation WIMS Microcontroller. . . . .	17
2.2 The WIMS Microprocessor architecture. . . . .	19
2.3 Measured energy for WIMS memory operations. . . . .	21
2.4 Simulation waveform from adder testing. . . . .	22
2.5 24-bit and 16-bit adder energy. . . . .	23
3.1 Leakage power for different memory access rates. . . . .	30
3.2 Access energies of memories in different process technologies. . . . .	31
3.3 Leakage trends for maximum and fixed access frequencies. . . . .	32
3.4 Drive current and leakage across processes. . . . .	33
3.5 Effect of high-K metal gate on gate leakage. . . . .	34
3.6 A die photo of the 180nm WIMS microcontroller. . . . .	36
3.7 Measured memory energy savings from executing code from the scratchpad memory instead of the main memory on the 65 nm WIMS microcontroller. . . .	37
3.8 Example program for explaining scratchpad management. . . . .	41
3.9 Comparison of energy savings from a scratchpad memory in 180, 65 and 32 nm technologies, using (a) the static allocation method and (b) the dynamic allocation method. . . . .	44
4.1 A schematic of the classical 2-stage opamp. . . . .	48
4.2 A schematic of a full rail-to-rail input stage with NMOS and PMOS input differential pairs in parallel. . . . .	49
4.3 Simplified schematic of class AB output stage. . . . .	50
4.4 Schematic of low-power opamp. . . . .	51

4.5	Transconductance of input differential pairs across common-mode input voltage.	51
4.6	Transistor-level schematic of low-power opamp. . . . .	52
4.7	Layout of the fabricated opamp. . . . .	56
4.8	Opamp bias voltage generation circuit. . . . .	57
5.1	Diagram of the conductivity sensor. . . . .	62
5.2	Figure of merit F, by ADC architecture. . . . .	65
5.3	Peak signal to quantization noise ratio (SQNR) by oversampling ratio (OSR). . . . .	68
5.4	Sigma-delta SNR versus gain of the integrating opamp. . . . .	68
5.5	Block Diagram of a Boser-Wooley 2 <sup>nd</sup> -order sigma-delta modulator. . . . .	69
5.6	Circuit schematic of the first stage of the sigma-delta modulator. . . . .	70
5.7	Complete simplified circuit schematic of 2 <sup>nd</sup> -order sigma-delta modulator. . . . .	72
5.8	Hogenhauer form of a sinc <sup>3</sup> filter. . . . .	74
5.9	Circuit simulation of modulator showing 88.6 dB SNR. . . . .	76
5.10	Frequency spectrum after the decimation filter. . . . .	76
5.11	Layout of the fabricated sigma-delta modulator. . . . .	77
5.12	Sigma-delta modulator measurement showing 50.0 dB SNR. . . . .	78
5.13	Sigma-delta modulator measured DC linearity plot. . . . .	79
5.14	Differential nonlinearity of ADC output. . . . .	79
5.15	Integral nonlinearity of ADC output. . . . .	80
6.1	Simplified schematic of 4-bit current scaling DAC. . . . .	84
6.2	Simplified schematic of a 2-bit voltage scaling DAC. . . . .	84
6.3	Simplified schematic of a 3-bit charge scale DAC. . . . .	85
6.4	Simplified schematic of a charge scaling DAC during clock phase p2. . . . .	86
6.5	Simplified schematic of extended resolution 6-bit charge scaling DAC. . . . .	87
6.6	Simplified schematic of an extended resolution 6-bit charge scaling DAC with a holding capacitor. . . . .	88
6.7	Timing diagram of DAC's four clocks during p2-p1 transition. . . . .	90
6.8	Simulated output of DAC. . . . .	91
6.9	Simulated differential nonlinearity of DAC. . . . .	91
6.10	Simulated integral nonlinearity of DAC. . . . .	92
6.11	Waveform from DAC simulation showing output glitch. . . . .	93
6.12	Layout of fabricated DAC. . . . .	94
6.13	Section of DAC output with ramp input. . . . .	95
6.14	Measured differential nonlinearity of DAC output depicted in Figure 6.13. . . . .	96
6.15	Measured integral nonlinearity of DAC output depicted in Figure 6.13. . . . .	96

7.1 Layout of the Utah Microcontroller. . . . .	99
7.2 Die photo of the Utah Microcontroller. . . . .	100
7.3 Size comparison of a PCB-based microsystem and the Utah Microcontroller. .	102

## ACKNOWLEDGMENTS

My parents nurtured my confidence throughout my life. I hope to be as loving of a parent as they each have been to me and my siblings. Even more than by teachings, they showed by example that true happiness is found through selflessness.

I have been fortunate to have wonderful teachers throughout my life. One inspirational figure was Dr. Eugene Clark, my high school physics and astronomy teacher. He showed a love of science that inspired many students beyond myself. I have had many other teachers like him throughout my undergraduate and graduate studies. Each of the members of my committee have helped me in unique and significant ways. It has been fascinating to see Dr. Zang lead his startup, Vaporsens, and I feel privileged to have been a part of it. Dr. Brunvand and Dr. Stevens taught me VLSI design, which is not only instrumental in this dissertation, but will likely be a focus of my career. Dr. Smith gave me personal assistance and encouragement with some difficult circuits, for which I am very grateful.

The work described in this dissertation builds strongly on the accomplishments of previous students. Matthew Guthaus created the original architecture of the WIMS Microprocessor, and it has been steadily improved by generations of students. Michael McCorquodale developed an excellent on-chip clock, Eric Marsman developed an impressive DSP block, and with Robert Senger went above and beyond in thoroughly documenting their work and helping ensure that it was passed on to the University of Utah. Keith Kraver's dissertation greatly assisted my own efforts to design a sigma-delta ADC. Nathaniel Gaskin, Spencer Kellis, and Jeff Campbell further improved the microprocessor by adding new instructions. Nathaniel and Spencer led our team through the University of Utah's first fabrication of the WIMS Microprocessor. Toren Monson helped develop the testing flow that enabled all of our publications on the microprocessor. I know there are many others whose contributions I am simply less familiar with. Dr. Brown has overseen the development of the WIMS Microprocessor from its conception, and it would never have reached the impressive state that it is in without him.

Beyond technical contributions, many collaborators have become close and trusted friends. Nathaniel Gaskin and Spencer Kellis were wonderful mentors as I began my

graduate studies. Rob, Amlan, Jeff, Toren, Shweta, Debdutta, and others were a joy to work with. Ondrej Novak's determination, hard work, and encouragement were vital in the final years of my program. I do not believe I would have graduated without him. Dr. Brown has been a kind and supportive advisor. His vast experience and knowledge make him an invaluable mentor.

And finally, I am especially grateful to my wife, for working every day to make life wonderful for our family.

# **CHAPTER 1**

## **APPLICATIONS OF MIXED-SIGNAL MICROSYSTEMS**

### **1.1 Introduction**

Device miniaturization and integration make in vivo medical data collection a tantalizing possibility for applications demanding microsystem sizes on the order of a few cubic millimeters. For example, a glaucoma monitoring device slightly larger than a single cubic millimeter can sample and record the internal pressure of the eye every 15 minutes, and transmit out the recorded data at the end of the day [1]. The Utah Microcontroller has been designed to act as a core component for these types of small low-power microsystems. The objective of the work described in this dissertation is to develop a single chip solution capable of supporting a variety of small low-power biomedical microsystems. To give concreteness to the design objectives, two specific target applications have been selected: a Smart Intravaginal Ring (S-IVR, described in Section 1.4) and an NO-releasing catheter (described in Section 1.5). Each of the circuits necessary to support these applications has been designed, fabricated, and tested, and all necessary components have been integrated into a single chip. While the complete microsystems have not been built, the integration and fabrication of a single chip with all circuitry components necessary to support the complete microsystems is a valuable and significant step forward.

Much of the work described in this dissertation is applicable to many other small low-power microsystems, and to the miniaturization of larger systems. For example, Dr. Ling Zang's research group is developing nanofibers capable of detecting the vapors of narcotics or explosives. The system being developed is much larger than the other microsystems described in this chapter, but it could be miniaturized for portable military use. The Utah Microcontroller, described in this dissertation, could be used in such a system. In this case, the portable system would have to perform significant signal processing while consuming very little power. The signal processing algorithms necessary for this application have been investigated and could be implemented in the Utah Microcontroller.

This chapter will describe the primary contributions of the dissertation, explore related microsystems, and describe each of the target applications. Subsequent chapters will describe the primary components designed by the author, analyze their advantages and suitability to the target applications, relate the target application requirements to specific circuit parameter requirements, and describe chosen circuit topologies and parameter optimization to meet circuit requirements.

## 1.2 Unique Contributions

The work described in this dissertation makes many unique contributions. A partial list of the most significant follows.

- Identification of common requirements for sensor applications: small size (preferably single chip), low power, analog sensor interface circuitry, and a wireless transmitter.
- The development of the Utah Microcontroller, a single chip that meets the requirements of many small low-power sensor applications, implanted and otherwise. Almost all commercial microcontrollers require additional circuitry chips to meet these sensor application requirements.
- The development of a highly-integrated circuitry chip, the Utah Microcontroller, incorporating a large variety of digital and analog circuitry, that meets the specific requirements of the S-IVR and NO-releasing catheter microsystems.
- An analysis showing that scratchpad memories are likely to continue to save power in the future, based on foundry transistor reports and measured data from the Utah Microcontroller.
- An analysis showing the cost of increasing addressable memory from 65 kB to 2 MB in a 16-bit microcontroller.
- A survey of commercial microcontrollers showing that a custom chip is necessary for a single circuitry chip to support the S-IVR and NO-releasing catheter.
- A survey of a large variety of ADC architectures and demonstration of how to select the best architecture for S-IVR and NO-releasing catheter requirements.
- A survey of a variety of DAC architectures and demonstration of how to select the best architecture for S-IVR and NO-releasing catheter requirements.

- A unique single-opamp switched-capacitor DAC architecture, and explanation of its benefits.

The target applications share the requirements of small physical size and low power consumption, and together represent a class of small low-power sensing systems. The desire to target this entire class of sensing systems contributes to the unique nature of the Utah Microcontroller.

In addition to the variety of circuits that needed to be developed, the design of the final integrated chip was undertaken by only two graduate students (Bennion Redd and Ondrej Novak), which motivated a desire for circuit reuse to reduce design time. Balancing many conflicting design requirements led to the unique design of the Utah Microcontroller.

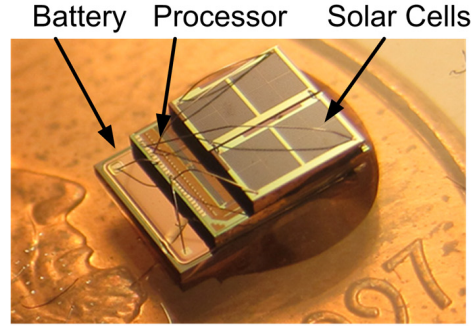
### 1.3 Related Microsystems

The small low-power microsystem targeted in this dissertation is similar to those used in a closely related research area: wireless sensor systems. These systems are built with low-power microprocessors, have integrated sensors, and communicate through a wireless network. They have been successfully used for healthcare [2] and environmental sensing [3]. However, most of these systems are built on printed circuit boards (PCBs), have many discrete components, and are far larger than the type of system proposed here. However, millimeter-scale sensing systems have been presented. One system—with a volume of only  $8.75 \text{ mm}^3$ —was targeted towards glaucoma monitoring and included both temperature and capacitive sensors [4], but that system did not include the chemical sensors necessary for many applications or analog circuitry optimized for micropotentiometric sensors. It is depicted in Figure 1.1.

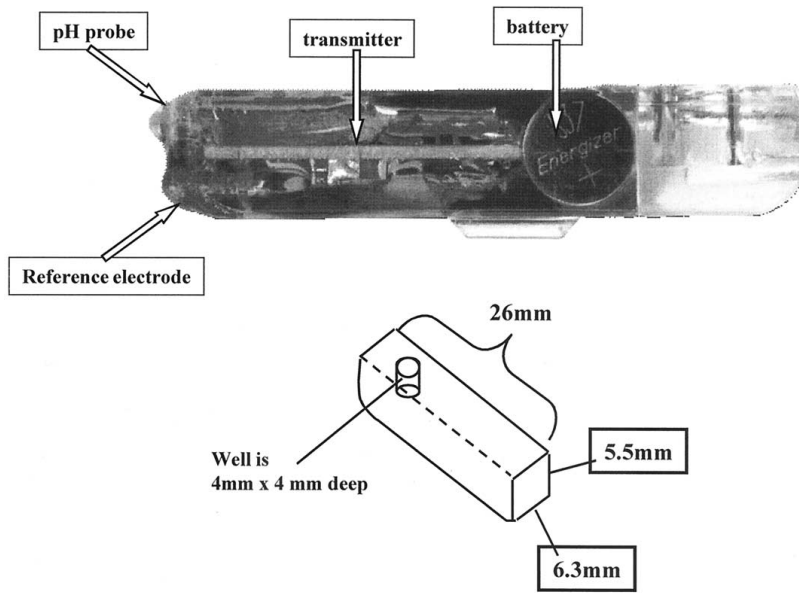
One commercially available system, depicted in Figure 1.2, is designed to test for gastroesophageal reflux disease [5]. This system is small, but not sufficiently small for the S-IVR. Typical IVRs have a 5 mm cross-sectional diameter. The S-IVR must match the size of an IVR, so the S-IVR's circuitry chip must fit in that 5 mm diameter. The system presented in [5] also does not have a conductivity sensor, even though research has shown that both impedance and pH monitoring are useful for gastroesophageal reflux diagnosis [6] [7]. A conductivity sensor will also be useful for the S-IVR, as discussed in Section 1.4.5.

Another wireless pH monitoring system [8] used an ISFET (ion-sensitive field effect transistor) for a pH sensor. The system was fabricated with only one chip in a commercial process. Another ISFET was fabricated to serve as a reference electrode, with an ion-





**Figure 1.1** A millimeter-size microsystem. Microsystem is on a penny, and contains a battery, processor, and solar cells. Image taken with permission from [4] ©2010 IEEE.



**Figure 1.2** A microsystem with a pH meter, transmitter, and battery. The system is  $26\text{ mm} \times 5.5\text{ mm} \times 6.3\text{ mm}$ . Reprinted by permission from Macmillan Publishers Ltd: The American Journal of Gastroenterology [5], ©2003.

blocking polymer deposited onto its surface. This polymer was “chosen to prevent hydrogen ions from reaching the pH-sensitive surface sites of the silicon nitride, while allowing the conduction of other ionic species” [8]. Because the drift characteristics of the reference electrode were different than those of the sensing electrode, the system would not be stable for the time frame required for the S-IVR. The pH sensor proposed for the S-IVR system will be provided by e-SENS. This commercial pH sensor uses a polymeric membrane on a micropotentiometric sensor and should have better drift characteristics. The RF transmitter was not integrated into the system-on-chip in [8]. The S-IVR system will also contain two chips (a sensor chip and the Utah Microcontroller), but will be smaller than [8]. The Utah

Microcontroller also consumes less power: less than 500  $\mu\text{A}$  while running at 10 MHz with a power supply of 0.8 V, while the system in [8] consumes 9 mA from a 3.3 V supply. Part of the advantage comes from the 65 nm process, as opposed to the 0.6  $\mu\text{m}$  process used in [8].

Although similar systems exist, there is not a wireless system that includes pH, temperature, and conductivity sensors packaged for the small size and harsh environment required for an S-IVR.

With the addition of a DAC, the same microcontroller can also be used to drive electrodes for the NO-releasing catheter. Available commercial microprocessors are surveyed in Section 2.1, which shows that there are no available chips with an integrated wireless interface and a DAC as required by these applications.

## 1.4 Smart Intravaginal Ring

A Smart Intravaginal Ring (S-IVR) microsystem is needed for testing microbicide IVRs designed to prevent transmission of AIDS. Such a microsystem could also be used for other purposes, such as fertility testing or user adherence studies. The proposed S-IVR system requires a microcontroller, interface circuitry, and sensors to suit the application. The design of most of this circuitry has been completed and integrated onto a single chip in an advanced process. A commercial sensor chip from e-SENS has the required sensors, and can be used for building a prototype Utah Integrated Microsystem for the S-IVR.

This prototype microsystem combines sensors, analog interface circuitry, and a full-featured microcontroller. It is a valuable and significant step toward a full microsystem capable of assisting in the test of microbicide IVRs to prevent the transmission of AIDS.

### 1.4.1 AIDS and the Promise of Microbicides

Although the annual number of new infections has declined over the last decade, there were still 2.3 million people newly infected with AIDS in 2012. AIDS is particularly a problem in Sub-Saharan Africa where the low status of women does not allow them to insist on use of a condom [9]. There are not yet any inconspicuous methods of protecting against the transmission of AIDS, but many research groups, including Dr. Patrick Kiser's lab at Northwestern University [10], have been working on new methods for preventing the transmission of AIDS: microbicide vaginal gels and Intravaginal Ring (IVR) technology. These technologies are very promising—they may eventually provide women with a private way of protecting themselves from AIDS.

A recent study by the Center for the AIDS Programme of Research in South Africa (CAPRISA) found that a microbicide gel reduced HIV infection by 39% [11]. While this

study shows that microbicides are effective, there are drawbacks to the gel form. The vaginal gels require daily application, which may lead to low user adherence. In the Carraguard vaginal gel clinical trial women adhered to the dosing regimen only 42% of the time. More intermittent dosage forms tend to have higher user adherence, and be more preferred by women. In one multinational clinical trial, after women had tried the Nuva-Ring®, 81% preferred it over a daily pill, primarily because of the convenience of an infrequent dosage form. A Brazilian study also indicated that women prefer an IVR over a gel partly because of the IVR's longer duration [12].

#### 1.4.2 User Adherence for Microbicide Trials

User adherence is a critical issue in assessing microbicide efficacy [13]. Self-reported usage data are often inaccurate or incomplete, so researchers have attempted other methods for assessing user adherence and activity. All of these methods have significant drawbacks. Biomarkers of semen such as semenogelin are only detectable in vaginal fluid for 3 days. Topical microbicides are detectable for weeks, but do not indicate timing of gel application or whether coital acts have occurred. Researchers sometimes request that study participants return products so that applicators or pills can be counted, but participants may forget or otherwise fail to return all unused products. The International Partnership for Microbicides (IPM) is developing a “smart applicator” that records when microbicide gel is applied [13], but no similar device exists to electronically record usage data of IVRs.

Usage data and coital activity without condom use could be accurately recorded by an S-IVR. The data collected by the S-IVR could be used in the testing of AIDS prevention devices, in user adherence studies involving birth control, or in medical fertility assistance.

#### 1.4.3 User Adherence Studies

Even though microbicides are an important research area, the S-IVR system has even broader applicability in user adherence studies. Birth control provides an excellent example of how dosage frequency and user adherence interact. Many oral contraceptives have a dose frequency of once per day. The Nuvaring®, an IVR contraceptive, has a dosage pattern of only once per month. Many professionals expected that Nuvaring would have a lower rate of unplanned pregnancy than the once per day oral contraceptives because user adherence would be improved due to the once per month dosage frequency. Unfortunately, Nuvaring appears to have the same rate of unplanned pregnancy as the oral contraceptives [14]. This suggests that user adherence is not as significantly improved by the low dosage frequency as was hoped. The reasons why could be better illuminated if an S-IVR was used to closely

monitor usage. For example, if users remove the IVR to clean it, then the impact of this cleaning on the IVR's effectiveness could be examined. Demographic information could be related to use profiles. Doctors could use this information to enhance user instructions and boost user adherence.

#### 1.4.4 Fertility Assistance

Data collected by an S-IVR can also be useful to the individual using the S-IVR: it can report valuable fertility information. Periods of maximum fertility can be predicted based on the conductivity [15] and temperature [16] information collected by the S-IVR. This information could be used by women attempting to conceive, or by fertility doctors in diagnosing fertility problems.

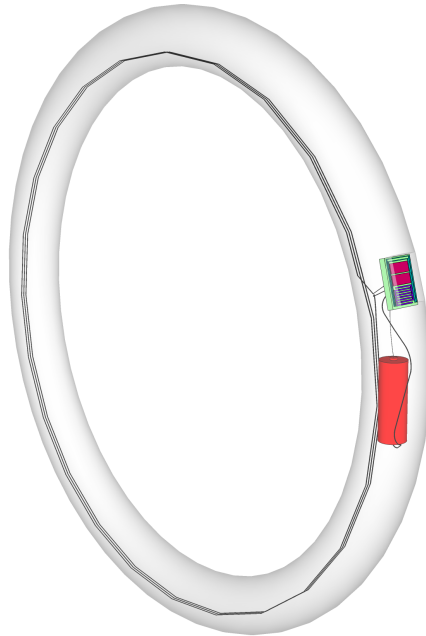
Using the S-IVR for fertility assistance would require system customization. The wireless interface would have to transmit data more frequently, and may need a different antenna to transmit through body tissue.

#### 1.4.5 Physical Description of S-IVR

The S-IVR, depicted in Figure 1.3, consists of a computerized microsystem which has had a silicone IVR molded around it. This microsystem reads data from three sensors: a pH sensor, a conductivity sensor, and a temperature sensor. The rationale for these three sensors is that a combination of these measurements can be used to distinguish when the IVR is in or out of the body, when it is being washed, and when it is exposed to seminal fluid. Measurement frequency is adjustable to maximize data collection for a given target lifetime. The processor analyzes the data to identify events of interest and stores only the data related to these events. This approach reduces the need for data storage by orders of magnitude. Since the system is completely programmable, different algorithms can be evaluated for event recognition.

The microcontroller is usually in a sleep state to save power, so that its battery can last through 28 days. At the conclusion of the 28-day study, the S-IVR is placed next to a wireless receiver to collect recorded data. The on-chip wireless transmitter uses ultra-wideband (UWB), and is described in Ondrej Novak's dissertation [17].

The entire microsystem must be sufficiently small to fit into an IVR with a cross-sectional diameter of only 5 mm. This small size is attainable because it has only two chips: a sensor chip, provided by e-SENS, and the Utah Microcontroller. The sensor chip is the larger of the two, and is 5.5 mm  $\times$  4 mm, which is slightly larger than the diameter of a typical S-IVR, but the slightly increased width probably will not be noticeable to most users. The



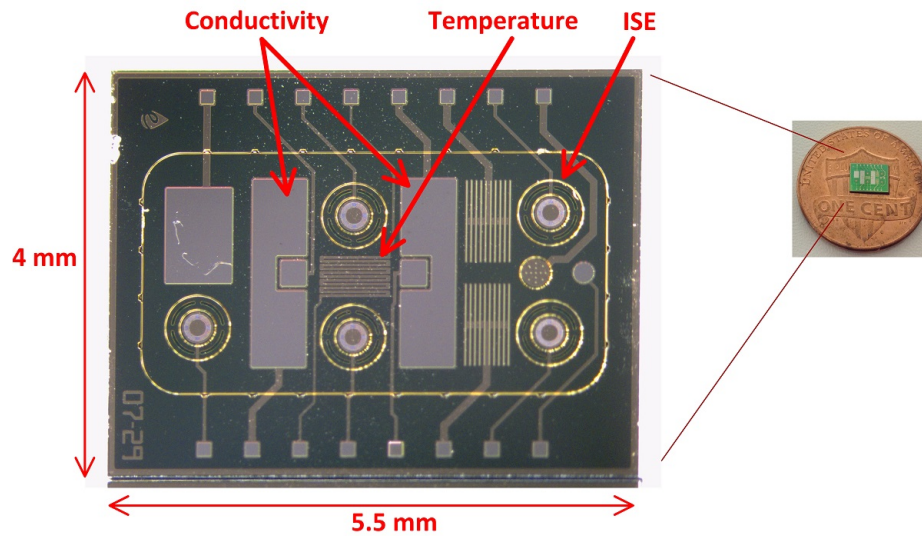
**Figure 1.3** Depiction of finished S-IVR. The encapsulating IVR is made transparent to show internal microsystem.

sensor chip is depicted in Figure 1.4. The microprocessor in the Utah Microcontroller was originally designed for the Wireless Integrated Microsystem Engineering Research Center (WIMS ERC) led by the University of Michigan, and is called the WIMS Microprocessor. The WIMS Microprocessor is combined with an integrated ADC, DAC, opamp (necessary to interface with solid-state chemical sensors), and wireless interface (to transmit out collected data) to form the Utah Microcontroller. The only components external to the two stacked chips are an antenna and a battery.

Although the entire system has not been built, a prototype of the circuitry chip has been fabricated and tested.

## 1.5 NO-releasing Catheter

More than 200,000 bloodstream infections occur each year in hospitals in the United States [18], most of which are related to catheters inserted into veins. Colonization of bacteria on the catheter is very common; in fact, Raad et al. found that bacterial colonization on central venous catheters was virtually universal [19]. Worse still, this near-universal bacterial colonization took the form of a biofilm. Bacteria form biofilms by producing a polymeric film that links the bacteria to each other and to the catheter, and bacteria in a biofilm can be highly resistant to antimicrobial treatment [20]. In the first 10 days, the



**Figure 1.4** Commercial sensor chip provided by e-SENS. Reused with permission ©2013 e-SENS.

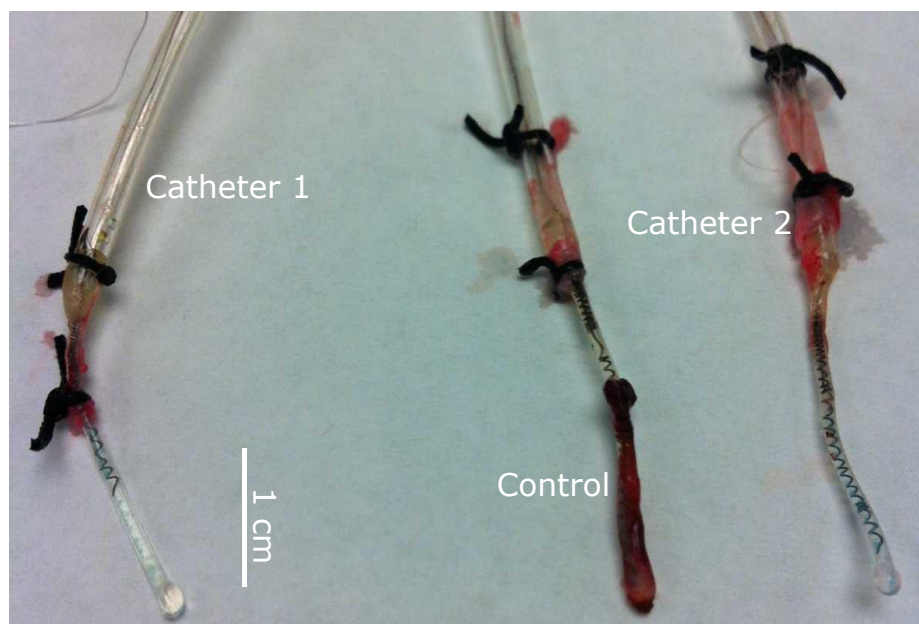
biofilm is primarily on the exterior of the catheter, but by 30 days after implantation, the biofilm primarily covers the catheter interior [19].

These bacteria come from either the patient's own skin, the healthcare personnel, or the fluid being delivered through the catheter if it has been contaminated, and their presence is difficult or impossible to prevent. Instead, many researchers have focused on how to prevent the bacteria from forming a biofilm on the catheter.

The release of nitric oxide (NO) can help to prevent biofilm formation on the catheter. Researchers have created small particles that release NO and deposited these particles on catheters. The particles were able to help prevent biofilm formation and reduce infection [21]. This prevention of biofilm is only one of several benefits to the release of NO from catheters [22]:

- Prevention of biofilm formation [21].
- Inhibition of platelet activation, which prevents blood clotting and the formation of scar tissue which could plug the catheter [23].
- Vasodilation, i.e., NO causes veins to increase in size, assisting delivery of fluid from the catheter [24].

An example of the dramatic reduction in blood clotting caused by the release of NO is shown in Figure 1.5. While NO-releasing particles have demonstrated the value of NO-releasing catheters, that approach has two serious drawbacks:



**Figure 1.5** Clotting prevention of NO-releasing catheters. Catheter 1 and 2 released an NO flux of 1.0 and  $0.9 \times 10^{-10} \text{ mol} \cdot \text{cm}^{-2} \text{min}^{-1}$  respectively, while the control catheter did not release NO [25]. Image provided courtesy of Dr. Mark E. Meyerhoff at the University of Michigan.

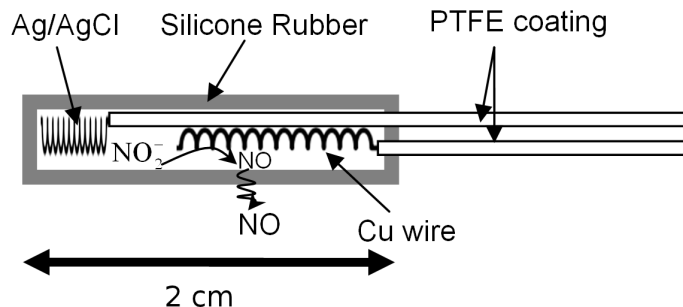
- The particles can only continuously release NO for 2–3 weeks, and cannot achieve longer-term release in a controlled manner.
- The base molecule, diazeniumdiolate, is unstable. This has made commercialization difficult [22].

Höfler et al. recently proposed another method of creating an NO-releasing catheter: the NO can be generated by an electrochemical reaction in a reservoir of sodium nitrite salt [22]. A conceptual prototype is depicted in Figure 1.6.

A positive pulse on the copper electrode will produce Cu(I) ions, which react with the sodium nitrite salt to produce  $\text{NO}_g$ , which can diffuse out through the catheter's walls. The catheter may be multilumen.

A naive first approach might be to merely connect a battery to the electrode, but a programmable microsystem is necessary for several reasons:

- The electrode must be cleaned with a negative pulse prior to the positive pulse to remove an oxide layer. Otherwise the electrode will not generate Cu(I) effectively. Even after the cleaning, Cu(I) can only be generated effectively for a short time.



**Figure 1.6** Diagram of NO-releasing catheter. The inner lumen to deliver fluid has been removed to simplify the drawing. Reproduced from [22] with permission of The Royal Society of Chemistry.

- Lengthening the time between pulses can make the reservoir of sodium nitrite salt last longer, so the catheter can be used for a longer period of time.
- The duration or frequency of the pulses may need to be varied over time, e.g., after 30 days, NO may need to be released more frequently to prevent biofilm formation.

The Utah Microcontroller can serve as the core of this microsystem. With a high current output DAC, it can drive the electrodes to the necessary voltages. The system could be made small and low power so that it could be implanted into an animal without the need for an external tether for a battery. It could potentially be reprogrammed wirelessly, to try a different pulse duration, frequency, or voltage. The conductivity sensor may even be able to detect biofilm formation on the catheter. These advantages can assist in researching this promising medical device or pursuing its commercialization.

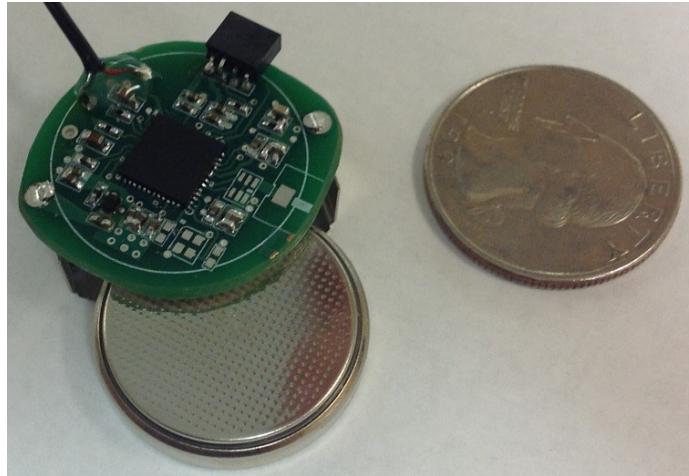
A single chip system will achieve the smallest size, and as shown in Section 2.1, there are no commercial chips available with both an integrated DAC and a wireless interface. A miniature system built on a printed circuit board (PCB) with commercial parts has already been developed by Ondrej Novak [26], and is depicted in Figure 1.7. The circuitry chip described in this dissertation will be a valuable step toward further miniaturizing that system.

## 1.6 Common Microsystem Requirements

This chapter gives an overview of many small low-power microsystems, particularly two microsystems selected to act as target microsystems for the Utah Microcontroller. These two applications share several requirements:

- Small size





**Figure 1.7** NO-releasing catheter microsystem. Developed on a PCB with commercial parts. Image taken with permission from [26] ©2013 IEEE.

- The S-IVR microsystem must fit inside a ring with a 5 mm diameter.
- The NO-releasing catheter microsystem is ideally only several cubic millimeters, to minimize impact on the movement of the test animal.
- Low power
  - The S-IVR must achieve a 28-day lifetime for a reasonable clinical trial to test microbicide IVRs.
  - The NO-releasing catheter must achieve a 28-day lifetime to test a prototype of a long-term implantable catheter.
- Wireless interface
  - The S-IVR requires a wireless interface to transmit out collected data at the conclusion of a study.
  - The NO-releasing catheter requires a wireless interface to transmit information such as battery life or biofilm measurements, which are useful in a study.
- ADC
  - The S-IVR requires an ADC with high input impedance for its pH sensor. The ADC is also necessary for conductivity and temperature sensors.
  - The NO-releasing catheter requires an ADC to monitor battery life and possibly biofilm formation on the catheter via a conductivity sensor.

- DAC
  - The S-IVR requires a DAC for its conductivity sensor.
  - The NO-releasing catheter requires a DAC with high output current to drive its copper electrode.

Just as these two systems share all of the above basic requirements, these same requirements satisfy a broad spectrum of similar low-power wireless sensor systems. The Utah Microcontroller is designed to satisfy all of these requirements in a single small chip. This dissertation will analyze the application requirements which drove the design decisions behind the Utah Microcontroller, and the results from testing after fabrication. The Utah Microcontroller is a single-chip solution to many small low-power sensor systems, which avoids having to design a custom chip for each specific application.

## CHAPTER 2

### UTAH MICROCONTROLLER

The Utah Microcontroller has been developed by Dr. Brown's research group for use in small low-power sensing microsystems. First the available commercial alternatives will be considered to establish the need for a custom solution, followed by a description of the Utah Microcontroller and the advantages of its core: the WIMS Microprocessor.

#### 2.1 Need for a Custom Microcontroller

It would be faster and cheaper to develop a system using an off-the-shelf commercial chip if one existed that could satisfy the application requirements established in Chapter 1. Therefore, the available commercial parts should be surveyed to determine if one exists that is suitable to those requirements.

There are a large number of microcontrollers available, with a variety of features and integrated peripherals. To survey these microcontrollers, searches of a component distributor website identified several companies with microcontroller offerings: Microchip Technology, Texas Instruments, Analog Devices, NXP, Maxim, and Silicon Laboratories. The microcontroller families from each of these companies were searched to find the product from each company that is most suitable to the applications described in Chapter 1.

Several microcontrollers have a number of the needed capabilities, but none have both these and the wireless interface. One example is the Silicon Laboratories C8051F396, which is part of their family of small size microcontrollers. This microcontroller meets many of the requirements found in Chapter 1. Packaged, it is only  $4\text{ mm} \times 4\text{ mm}$ , and is available as an unpackaged bare die which measures only  $2\text{ mm} \times 2\text{ mm}$ . It includes 10-bit current DACs capable of driving up to 2 mA of current at 3 V. The DACs only consume 60  $\mu\text{A}$  of current beyond the full-scale output current. A 16-channel 10-bit ADC is also included. The ADC draws up to 1 mA of current, but can be used intermittently to conserve energy. The MUX input to the ADC has an input impedance of 1.6 k $\Omega$ , which is insufficient to read data from a micropotentiometric sensor. The microprocessor draws only 200 nA when sleeping

and 250  $\mu\text{A}$  when operating from the 80 kHz oscillator. These specifications indicate it is very close to meeting the target application requirements, but the microcontroller lacks an important component: the wireless transmitter. Without this component, collecting data at the end of a study involving the S-IVR, for example, would be very difficult.

The requirement for an integrated wireless interface narrows the field considerably. For example, Silicon Labs' Parametric Search lists 842 configurations of microcontrollers available [27], but their Wireless Solutions Guide only lists 28 wireless microcontrollers [28], categorized into 6 families on their webpage [29]. For brevity, all other microcontrollers considered below will be limited to those with an integrated wireless interface.

For most chip manufacturers, even the smallest and lowest-power chips were too large and drew too much power for our target applications. For example, of all the NXP chips, the JN516x family is the smallest and lowest power chip that integrates a wireless interface and an ADC. Its RF interface uses the ZigBee standard, and it includes a 4-channel 10-bit ADC, but the smallest module is 16 mm  $\times$  21 mm, which is far too large for our target applications.

Another example is the Microchip rPIC12F675, the lowest-power wireless PIC microcontroller with an ADC. It typically draws a very low 18  $\mu\text{A}$  typical at 3 V when running off the low-power 32 kHz oscillator, but the package is 8 mm  $\times$  7 mm, and there is no DAC [30].

The most suitable Silicon Laboratories chip for our target applications is the Si106x family, because it is the smallest of the microcontrollers with wireless capability and an integrated 10-bit ADC. This family supports input voltages of 0.9 – 3.6 V via its DC-DC converter. It is relatively low-power, drawing just above 90  $\mu\text{A}$  at 32 kHz. Unfortunately the packaged chip is above application requirements, at 5 mm  $\times$  6 mm, and does not include an integrated DAC or an opamp [31]. None of Silicon Labs' chips include both an integrated wireless interface and a DAC [28].

One way to reduce power requirements further is to use microchips developed to support near field communication (NFC). One example is Texas Instruments' (TI) RF430FRL15xH, which incorporates TI's common low-power microprocessor, the MSP430. This packaged chip is only 4 mm  $\times$  4 mm, meeting size requirements. An NFC antenna can be connected to the system to read data from it, and it has a 14-bit sigma-delta ADC and an integrated temperature sensor. This is the only NFC chip that has an integrated, programmable microcontroller, and does not require a separate chip. It has 2 kB of FRAM, which may be insufficient for data collection for the S-IVR, and does not include a DAC [32].

The CC430 is the wireless integrated version of Texas Instruments' MSP430 microcontroller [33]. The MSP430 is a very common microcontroller for low-power applications.

The CC430F5137 is the version most suited to our target applications due to its integrated wireless interface and ADC. It consumes only  $1\text{ }\mu\text{A}$  in sleep mode. However, the package size is  $7\text{ mm} \times 7\text{ mm}$ , which is still too large and there is no integrated DAC [34]. TI offers an integrated wireless chip based on the 8051 microcontroller, the CC1110, which consumes only  $0.5\text{ }\mu\text{A}$  in sleep mode and fits into a smaller  $6\text{ mm} \times 6\text{ mm}$  package, but this is also too large for the S-IVR and lacks an integrated DAC [35].

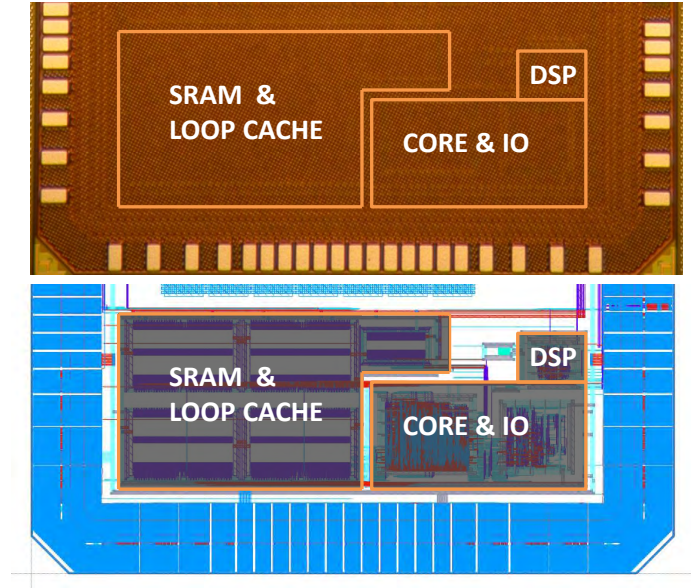
A list of which microcontrollers met the important requirements listed in Chapter 1 is in Table 2.1. In fact, of the manufacturers listed in Section 2.1, none of them carry a microcontroller with both an integrated wireless interface and a DAC. The only way to satisfy all application requirements with a single-chip solution is to employ a custom chip design.

## 2.2 Background on the WIMS Microprocessor

The microprocessor in the Utah Microcontroller is called the WIMS Microprocessor because it was built for the Wireless Integrated Microsystem Engineering Research Center (WIMS ERC) at the University of Michigan. The author’s research group has already developed several microsystems built around the WIMS microprocessor. The first of these systems was configured as a remote environmental sensor with an analog front-end to interface with solid-state sensors [37]. The second version, aimed at cochlear prosthesis applications, included a DSP block performing the continuous interleaved sampling (CIS) algorithm, which is used to translate sound into signals to drive the electrodes implanted in the cochlea [38]. These two systems were fabricated in TSMC  $0.18\text{ }\mu\text{m}$  CMOS. An adaptation for use as a neural recording device was fabricated in IBM’s  $65\text{ nm}$  technology [39], and is depicted in Figure 2.1. It is referred to as the third-generation WIMS Microcontroller. Measurements of the third-generation WIMS Microcontroller show that it achieves very low

**Table 2.1** Commercial microcontrollers and application requirements. Application requirements are discussed in detail in Chapter 1.

Microcontroller	Low Power	Small Size	ADC	DAC	RF
Silicon Labs C8051F396	✓	✓	✓	✓	
Silicon Labs Si106x	✓	✓	✓		✓
Microchip rfPIC12F675	✓		✓		✓
TI RF430FRL15xH	✓	✓	✓		✓
TI CC430F5137	✓		✓		✓
TI CC1110	✓		✓		✓



**Figure 2.1** Layout of the third-generation WIMS Microcontroller. This is the immediate predecessor to the Utah Microcontroller. The die photo on top shows metallization with block outlines added. The lower image is a screenshot of the digital logic layout. Image taken from [36] ©2012 IEEE.

power consumption. The newest version, the Utah Microcontroller, fabricated in Global Foundries' 65 nm process, also contains the WIMS Microprocessor and integrates a variety of analog circuitry.

The motivation for using a 65 nm technology is that it is lower power and results in a smaller chip size than is possible in 0.18  $\mu\text{m}$  CMOS. Fortunately the MOSIS Educational Program mitigated cost concerns for this project [40]. In a commercial implementation of sensor microsystem, cost would be a significant consideration in the choice of process node.

The third-generation WIMS Microcontroller has been tested, and properly executes all instructions in its ISA. Interrupts, block transfer instructions, and other features are fully functional. The core power draw while executing code is less than 500  $\mu\text{A}$  while running at 10 MHz with a power supply of 0.8 V. This low power draw indicates that with the addition of an effective sleep mode, the power targets necessary for the S-IVR are obtainable. The Microcontroller benefits from architectural features developed in earlier versions that support very low-power operation, such as an innovative scratchpad memory [41] and block transfer instructions [39]. These architectural features have enabled the author and his colleagues to examine the address space architecture of the WIMS Microprocessor [36] and study how process scaling affects the energy savings of scratchpad memories [42]. These studies have been incorporated into this dissertation in Section 2.3 and Chapter 3.

The Utah Microcontroller focuses on recording and processing data from chemical and electrical sensors. It integrates the WIMS Microprocessor with analog circuitry, such as an opamp for buffering signals, an ADC, DAC, voltage regulators, sleep mode, and UWB wireless interface onto a single chip. The ADC will be covered in Chapter 5, the DAC in Chapter 6, and the opamp in Chapter 4. The voltage regulators and UWB interface were designed by Ondrej Novak, and are described in his dissertation.

## 2.3 Architecture for Increased Address Space in an Ultra-Low-Power Microprocessor

Section 2.3 is derived from [36] ©2012 IEEE, and estimates the costs required to expand the address space of the WIMS Microprocessor to 24 bits instead of 16 bits.

### 2.3.1 Introduction

Many small low-power sensing systems have significant data storage requirements. For example, the number of channels and resolution of implantable neural recording systems continue to increase. A neural prosthetic would have to include memory for recording neural activity, as well as complex programs to analyze and interpret that neural activity.

One architectural challenge for low-power sensing systems is to provide an address space large enough to store all collected data while maintaining the low-power benefits of a small instruction size. For example, a 16-bit microprocessor can achieve lower energy per instruction than a 32-bit microprocessor, but its native address space is only 64 kB. A 32-bit microprocessor can address a larger memory, but it uses more power for each executed instruction and memory access.

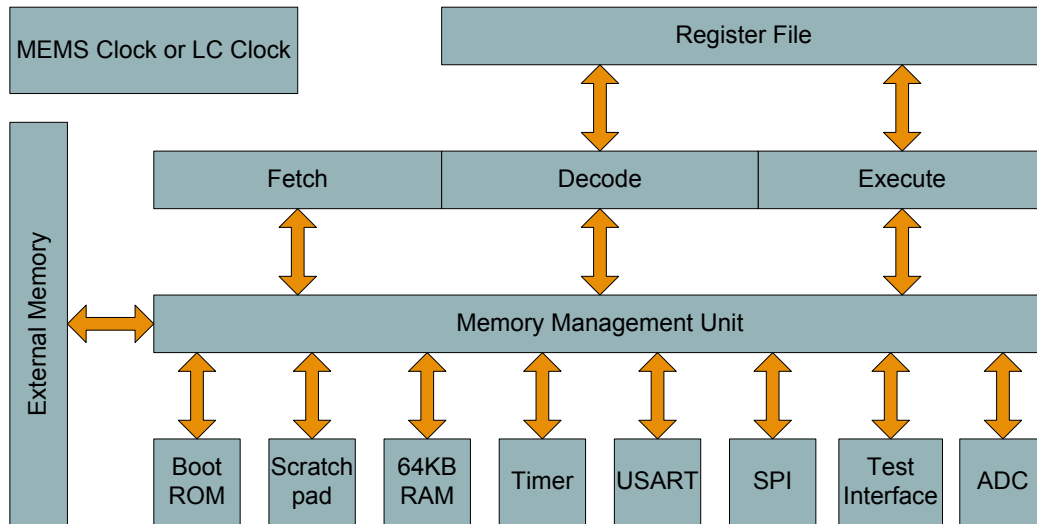
Virtual memory systems and dedicated memory controllers, while able to increase addressable memory, consume enormous amounts of power in the context of sensor system microprocessors. Often, such systems require off-chip memory which increases I/O power. Three-dimensional wafer bonding may soon provide a method of fabricating the circuitry and memory in separate optimized processes, without incurring the penalty of additional I/O power. In this case, a memory controller can be eliminated if the ability to address a larger amount of memory is integrated directly into the pipeline.

Section 2.3 describes a hybrid solution to satisfy the conflicting requirements of low power and large address space: keeping instruction and data register widths at 16 bits, while increasing the size of address registers to 24 bits. The WIMS Microprocessor, which implements this approach, has been fabricated in a 65 nm IBM CMOS process. Using measured data from a fabricated chip, the efficiency of this solution is evaluated in terms of

the energy and area requirements needed to implement the larger address space relative to a hypothetical 16-bit version of the WIMS Microprocessor. This hybrid solution harmonizes the advantages of the lower execution power of a 16-bit data path and a large address space, but eliminates the need for a memory controller to translate addresses between different address spaces.

### 2.3.1.1 WIMS Support for 24-bit Addresses

The WIMS microarchitecture is depicted in Figure 2.2. The 24-bit address width enables it to directly address 2 MB, while the execution pipeline keeps operand and instruction sizes at 16 bits. The execution pipeline consists of three stages: fetch, decode, and execute. The buses between these stages and the memory management unit or register files are depicted as arrows in Figure 2.2. The addresses passed from the pipeline to the memory management unit were increased to 24-bits. Several state registers in the pipeline such as the PC were also extended to 24-bits. The register file is sixteen 24-bit address registers that are byte writable, and sixteen 16-bit data registers. The memory address adder in the decode stage was increased to 24-bits. The memory management unit decodes the 24-bit addresses and directs the request to the appropriate peripheral. The memory management unit varies in size depending on the number of addressable peripherals included in the system. In the 65 nm implementation, the memory has been implemented in four small banks of 8 kB and a small scratchpad memory [41] [42] to reduce power when execution is dominated by a small loop.



**Figure 2.2** The WIMS Microprocessor architecture. Image taken from [36] ©2012 IEEE.



## 2.3.2 Measurements

### 2.3.2.1 Leakage

The WIMS Microprocessor has sixteen 24-bit registers: 14 address registers, the Program Counter (PC), and a Stack Pointer (SP). Some of these values are also passed along pipeline stages, so 24 registers are increased to 24-bits for this expansion. Therefore, 192 additional flip-flops were added to support 24-bit addressing. The total area expansion due to the additional flip-flops is  $1,555 \mu\text{m}^2$  or 2.5% of the execution pipeline area. A group of cells were randomly selected from the layout, and their leakage energy per area was calculated from the cell library datasheet. The flip-flop cell has leakage energy per area 20% higher than the average of the randomly selected cells, so the total additional leakage energy is calculated to be approximately 3%.

### 2.3.2.2 Per-Instruction Energy Measurement

The 24-bit registers are stored as three byte-addressable 8-bit registers. The additional energy of accessing a 24-bit register instead of a 16-bit register can be estimated by comparing a pair of instructions that differ only by asserting a read or write signal for one byte of the register.

To calculate the energy consumed per instruction, an instruction was placed inside a loop 400 cycles long. The current was measured 10 times for each test case. The average energy per instruction ( $EPI$ ) was found using this equation:

$$EPI = (V \cdot I_{avg} \cdot CPI) / F \quad (2.1)$$

$V$  is voltage,  $I_{avg}$  is average current,  $F$  is frequency, and  $CPI$  is cycles per instruction. The  $EPI$  was found for two similar instructions, which differed only by one instruction performing an additional operation, and the energy of the additional operation was estimated as the difference in energy between the two instructions. The measurement is only approximate because of additional switching such as instruction decoding that cannot be fully isolated. NOOP (“no-operation”) instructions were interleaved between instructions in the loop to help isolate the desired operation. The microprocessor was run at 120 MHz at 1 V. Data on how energy consumption for the WIMS Microprocessor varies with frequency and voltage have been published previously [39].

### 2.3.2.3 Read and Write Energy

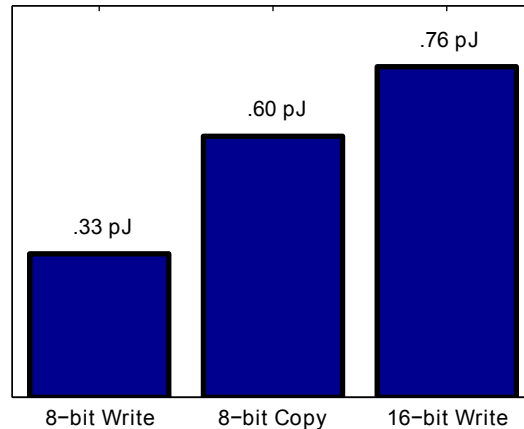
To estimate the additional energy to read and write a 24-bit register instead of a 16-bit register, two copy instructions were compared: CAR and CDR. CAR, “Copy Address

Register,” copies a 24-bit address register to another 24-bit address register. CDR, “Copy Data Register,” copies a value in a 16-bit data register to another 16-bit data register. The measured energy difference between CAR and CDR was 0.59 pJ, or 1.3% of the active energy of one clock cycle.

The 8-bit write energy was estimated by comparing the CDAR and CDARB instructions. The CDAR instruction copies 16 bits from a data register to an address register. CDARB is identical, but only asserts the write signal for 8 bits. For the comparison, both CDAR and CDARB alternated between writing 0x0000 and 0xffff. The difference between them is approximately the energy to write 8 bits of the opposite value to the stored bits, and was 0.33 pJ, or 0.7% of active energy.

The 16-bit write energy was also estimated. The MUL instruction multiplies two 16-bit values, and writes the 32-bit result to two data registers. The MULS instruction is identical, except that only the upper 16-bits are written. The full multiply is performed in both cases. The difference in energy between these two instructions is the 16-bit write energy, and was 0.76 pJ, or 1.7% of the active energy of one cycle. The compared instructions use multiple cycles to complete, so the comparison to the active energy of one cycle is only provided for context.

The numbers reported are graphed in Figure 2.3. These measurements assumed maximum bit-flipping, whereas during normal operation the upper bits of most of the state and address registers in the system would flip only rarely.



**Figure 2.3** Measured energy for WIMS memory operations. Energy of each of the listed operations was measured with maximum bit flipping. Image taken from [36] ©2012 IEEE.

### 2.3.2.4 Adder Energy

The adder in the instruction decode stage is 24-bits wide instead of 16-bits. By default, the PC is fed into the 24-bit adder. If the microprocessor is primarily using a small part of the 24-bit address space, then the upper bits fed into the 24-bit adder will not usually change from cycle to cycle. This section illustrates that the active energy consumed by the adder depends primarily on the number of bit flips in an addition, rather than on the total width of the adder. The increase in active energy for increasing bit flips will be measured for both a 16-bit and 24-bit adder.

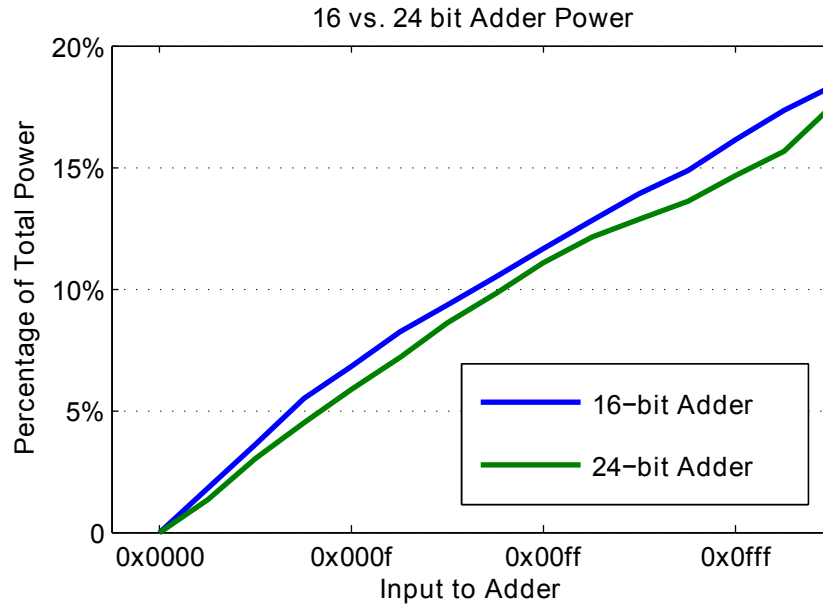
The microprocessor's decode stage has a 24-bit adder, but the ALU has a 16-bit adder, so the energy consumed by each adder can be estimated independently and compared. In these tests, the 16-bit adder was operated with the ADD instruction, interleaved with NOOPs so that the inputs to the adder were reset to 0x0000 between each ADD operation. The output of the adder was written to one of the source registers, and the value was allowed to increment continually. Simulation waveforms from the Verilog description of the microprocessor (similar to the one in Figure 2.4) were examined to confirm that signals associated with significant power draw (such as the inputs to the 24-bit adder) were stable while the 16-bit adder's energy was being measured.

To measure the energy used by the 24-bit adder, a "Load Byte with Register Offset and Update Address Register" (LDBU) instruction was used. This instruction adds the values of a 24-bit address register and 16-bit data register to calculate a memory address, loads a byte from that address, and then updates the address register to the new address. The NOOP instruction feeds the PC to the 24-bit adder, so LDBU instructions referring to zeroed registers were interleaved to reset the 24-bit adder inputs to zeroes. Waveforms such as the one depicted in Figure 2.4 were examined in a Verilog model of the microprocessor to verify that the inputs to the 16-bit adder did not change while the 24-bit adder energy was measured.

The number of 1's fed to the adder was increased from the LSB to the MSB (e.g.,



**Figure 2.4** Simulation waveform from adder testing. Waveform illustrates that the inputs to the 24-bit adder are reset every other cycle, and the inputs to the 16-bit adder remain static during the testing of the 24-bit adder. Image taken from [36] ©2012 IEEE.



**Figure 2.5** 24-bit and 16-bit adder energy. Increase in energy per cycle to add numbers to a running total, measured as a percentage of total active power. Image taken from [36] ©2012 IEEE.

0x0001, 0x0003, 0x0007, 0x000f, etc.) and for each input number, the power was measured while that number was repeatedly added to a running total. These values were selected to slowly increase the number of bit flips that occur with each addition. The running total was periodically reset to prevent overflow. The results, depicted in Figure 2.5, indicate that the marginal active energy consumed by the 16-bit adder is actually larger than that consumed by the 24-bit adder. This erroneous result is due to a small measurement error: the 24-bit adder must be reset every 8 cycles to keep memory addresses valid, and during the reset a small non-zero value is fed into the 24-bit adder. Therefore, the difference between the 0x0001 and 0x0003 cases is less significant than with the 16-bit measurement. However, the results illustrate that if upper bits seldom flip, the use of a larger adder does not increase the total switching power relative to static inputs.

### 2.3.3 Minimizing Code Size

Power due to instruction bit switching is minimized by keeping the instruction size at 16-bits (except for a few two-word instructions), but the number of addressable registers is limited by the number of bits used in the instruction to specify a register. The WIMS architecture implements a windowed register scheme in order to minimize code size while allowing a reasonable number of address registers [43]. Only 3 bits of most instructions are used to specify one of 8 registers available in the currently-selected window. The window

can be changed by another instruction. A few unique instructions can access any register or copy registers across windows.

The WIMS ISA also offers a number of addressing options to allow 16-bit instructions to load and store values in a 24-bit address space and registers. Load and store instructions can specify an 8-bit immediate offset, so one address register can be used to specify any address in a 256 B block. There are 12 address registers, 6 of which are active at a given time, and each can be used to point to a different block in memory. If a 16-bit data register is used to store the offset value, the block can be expanded to 64 kB. If a block size larger than this is required, the upper 8 bits of the address register can be loaded as an immediate with a single instruction (LDABI). These architectural features mean that most sequential loading and storing operations take only one or two instructions.

Jump and branch instructions are similar to the case for load and store. A relative jump, branch, or subroutine call can go forward or backward up to 127 words. An absolute jump or subroutine call (a two-word instruction) can jump to anywhere in memory. Even a very large program would not require very many extra cycles, as a single extra word is sufficient to jump to anywhere in addressable memory.

### 2.3.4 Alternatives

One alternative to our method is to fix the width of both the address and data paths, e.g., implement a 24- or 32-bit system. A system with larger instructions would result in higher power because of additional switching energy to accommodate the longer instruction size and more complicated instruction decoding. Such systems usually target high-performance applications rather than low-power applications, where higher performance requirements typically correspond to larger area and power budgets.

Researchers have proposed supporting multiple sizes of instructions [44] so a smaller execution path can be used whenever possible. This approach requires a larger die area and more hardware overhead than our approach, but provides more flexibility. Another option is to clock-gate unused bits in execution units [45]. Neither of these approaches is targeted at increasing the address space available to a low-power processor, but focus instead on reducing the power consumption or code size of a processor that already has a large address space.

A memory hierarchy approach, common in high-performance systems, requires data to be copied either to a smaller local memory or a cache before they are accessed or executed. This is a good approach combining a slow, large, low-power memory with a small and fast memory, but requires more copying and more complex hardware. Our approach assumes

performance requirements are lax enough to access low-power on-die memory directly. This situation may become more common as wafer-bonding technologies allow a low-power logic process to be combined with a memory-optimized process.

### 2.3.5 Conclusion

Section 2.3 has quantified the energy and area requirements to obtain a 24-bit address space with a 16-bit data path. The additional registers increased pipeline area by 2.5%, and leakage energy by approximately 3%. The active power impact cannot be measured without a specific benchmark, but writing an additional 8 bits, with all 8 bits flipping, to a single register increases active energy by only 0.33 pJ, or 0.7% of the total active energy per cycle. Although many pipeline registers have been expanded to 24 bits, only a few (if any) of the added 8 upper bits switch in a given cycle. The incorporation of 16 address registers and several addressing schemes such as immediate-offset and register-offset helps minimize the code expansion caused by the small instruction size. Thus, with a creative combination of addressing options, the WIMS architecture exemplifies an efficient and flexible method for achieving a large memory space with a low-power microprocessor. The address space has been increased by 256-fold, from 64 kB to 2 MB with only a 2.5% increase in area, 3% increase in leakage, and a similarly small percent increase in active power.

## CHAPTER 3

### SCRATCHPAD MEMORY

This chapter is derived from a journal paper published in the *Journal of Low Power Electronics and Applications* [46]. It focuses on a low-power feature used in the WIMS Microprocessor: a scratchpad memory.

#### 3.1 Introduction

Modern microprocessors may use techniques, such as small low-power scratchpad memories or very small instruction caches, to reduce power consumption. Both of these types of memory reduce the average power of on-chip memory accesses by making the most-frequent accesses to these small memories. The scratchpad memory uses even less energy than a cache, because it does not have tag storage or comparison circuits. These techniques have been demonstrated to reduce power consumption in older technologies, but their benefits need to be reevaluated for new nanometer-scale processes. Higher leakage power in these processes could outweigh the energy savings gained by adding a scratchpad memory, so the assumptions made in previous analyses need to be reexamined.

This paper describes a model that quantifies the benefits of scratchpad memories using memory access rates, access energies and leakage power data. The model is used, along with data from CACTI (a cache and memory access time model) simulations, to show how both process technology and application characteristics impact the energy savings resulting from the inclusion of a scratchpad memory. The concepts demonstrated by this model put the importance of leakage power into perspective. In this expanded version of our earlier analysis [42], leakage trends are evaluated using Intel's process technology reports to see whether scratchpad memories will be viable power-saving architectural features in the foreseeable future.

Measured data from the WIMS (Wireless Integrated MicroSystems) microcontroller [39] can be used to verify this conclusion, because this microcontroller has a scratchpad memory and has been implemented in both 180 and 65 nm processes. The energy profile of a 32

nm version of the microcontroller can be projected using reported Intel transistor data. Energy savings due to the incorporation of a scratchpad memory will be projected for several benchmark programs using the previously described model and measured data from the WIMS microcontroller.

## 3.2 Background

Scratchpad memories are small, low-power memories that can reduce power consumption by providing low-energy access to frequently used data or program fragments, such as loops. These memories are most effective for applications that are dominated by many small loops or that have small data blocks that are frequently accessed [47].

One precursor to the WIMS approach to scratchpad memories was the employment of a very small direct-mapped instruction cache, called a filter cache [48]. It reduced access energy, but was still subject to thrashing and conflicts. This filter cache was refined into a loop cache (LC) by using information from a profiler to allow the compiler to place instructions into the loop cache optimally. An instruction was added to inform the processor of which blocks should be placed into the loop cache [49]. Other approaches recognized the loop to be placed into the loop cache from a new “short backward branch” instruction, avoiding the need for profiling or tag comparison circuitry in the LC [50, 51].

To overcome the burden of added hardware complexity in the loop cache, a scratchpad memory (a memory without a hardware controller) was developed. In this approach, the compiler chooses frequently used data, or code fragments [47], or virtual memory pages [52] to statically reside in the scratchpad memory. These objects are accessed by an absolute memory address and are software managed, which reduces hardware complexity by eliminating both the tag comparison circuitry and the LC controller. When the most frequently accessed objects are placed in the scratchpad memory, less energy is used by the system than would have been used with either the filter cache or LC approaches [53]. Profiling is used in this case.

The WIMS microcontroller compares two scratchpad memory approaches: a static approach, in which code or data blocks in the scratchpad memory remain constant for a compiled program; and a dynamic approach, in which the contents of the scratchpad memory change over the course of the execution of a program [41]. These approaches are explained in more detail in Section 3.6.3.



### 3.3 Motivation

Early research on scratchpad memories was conducted on prenanometer processes [48–51, 53]. As semiconductor processes are scaled, feature size reduction and material changes affect power consumption and memory performance, which could affect the benefits of scratchpad memory. In particular, increased leakage power could outweigh the dynamic power savings of scratchpad memories. Researchers have anticipated [54] that as processes are scaled, leakage will become a greater problem.

Recent research on scratchpad memory has addressed the increasing importance of leakage power by finding ways to minimize leakage. Guangyu et al. [55] split the scratchpad memory into different banks and considered turning off one or more banks for some loop nests to minimize leakage energy. Each loop nest was analyzed using integer linear programming to determine the optimal number of scratchpad memory banks to use. Kandemir et al. [56] mapped data with similar temporal locality to the same scratchpad memory bank to maximize bank idleness and to increase the chances of that bank entering a drowsy state, thereby decreasing power. Huangfu et al. added compiler instructions, so that their compiler can activate and deactivate groups of words in the scratchpad memory to significantly reduce leakage without a significant performance penalty [57]. Takase et al. [58] used data on leakage trends from CACTI 5.0 [59] to consider how leakage trends impact scratchpad memories across two process transitions.

The study described in [42] added measured data from the WIMS microcontroller implemented in two process generations, another simulated process transition (for a total of three) and a model to assist in characterizing measured data from a fabricated processor. The present paper expands upon that study by examining more recent data on leakage trends from industry transistor reports from SoC processes, energy savings projections for the WIMS microcontroller based on the most recent transistor reports and information on the more effective dynamic scratchpad memory allocation algorithm, which was implemented for the WIMS microcontroller.

## 3.4 Conceptual Framework

### 3.4.1 Models

To clarify the relationship between energy savings, access rates (the percentage of total accesses that are directed to the scratchpad memory) and access energies, a simple mathematical model was developed. Dynamic energy savings ( $ES$ ) are calculated from scratchpad access rates ( $AR$ ), scratchpad fetch energy ( $SFE$ ) and the main memory fetch energy ( $MFE$ ):

$$1 - ES = \frac{(AR \cdot SFE + (1 - AR) \cdot MFE)}{MFE} \quad (3.1)$$

This model applies to both systems with on-chip and off-chip main memories. The denominator of the fraction on the right-hand side of (3.1) represents energy without a scratchpad memory, and the numerator represents energy usage with a scratchpad memory. Simulations presented in [41] for the 180 nm process assumed the model described by (3.1).

To account for leakage energy, (3.1) was updated to include the two new variables  $SLE$  (scratchpad leakage energy) and  $MLE$  (main memory leakage energy):

$$1 - ES = \frac{(AR \cdot SFE + (1 - AR) \cdot MFE) + MLE + SLE}{MFE + MLE} \quad (3.2)$$

$SLE$  and  $MLE$  are defined to be the average leakage energy per memory access, due to the scratchpad memory and main memory, respectively. This equation assumes that the scratchpad and main memory latencies are equal, which is pessimistic for many systems, but accurate for the WIMS microcontroller, which has an on-chip main memory. This model predicts that energy savings will be positive any time the numerator on the right-hand side of the equation is smaller than the denominator. This condition, along with algebraic simplification, results in (3.3), which indicates whether the scratchpad memory saves power:

$$SLE < AR \cdot (MFE - SFE) \quad (3.3)$$

These equations will be used with data generated by the CACTI 5.3 Pure RAM Interface [59] to demonstrate how to evaluate the benefits of the scratchpad memory and to show which scaling trends can be used to predict the future benefit of a scratchpad memory. The CACTI-generated SRAMs used to generate the figures in Section 3.4 have one bank, one read/write port and no other ports and read out 16 bits. These memories are optimized for low standby current. In Section 3.6, these equations will use measured data from the WIMS microcontroller to project energy savings for several benchmarks.

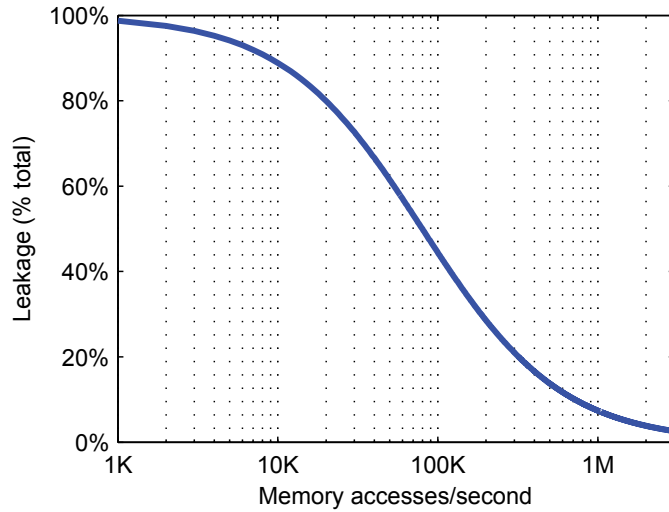
### 3.4.2 Relative Importance of Active and Leakage Energy

A small memory, such as a scratchpad memory, will usually have lower access energy than a large memory, like the main memory. On the other hand, the addition of a scratchpad memory will increase memory leakage power. If the scratchpad memory is used to reduce power, the reduction in access energy must outweigh the increase in leakage power. The scratchpad memory is usually very small compared to the main memory, so the increase in leakage power is generally very small, as well, but may be larger in future processes [54,60]. The relative importance of the access energy compared to the leakage energy depends on

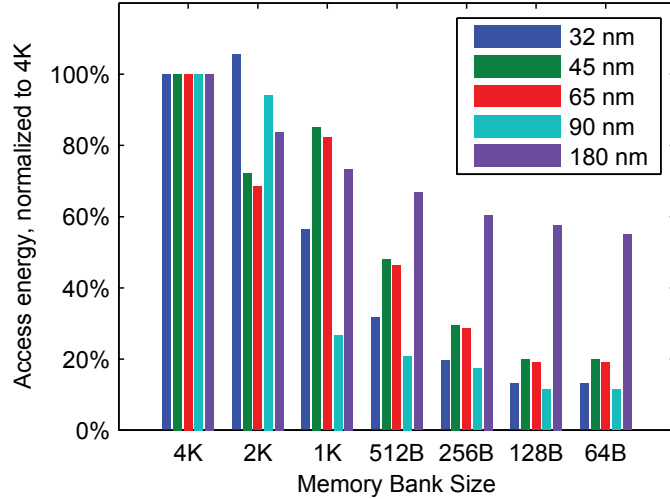
how frequently the memory is accessed in a given application. This concept is illustrated by Figure 3.1, which uses data from CACTI 5.3 to show leakage power as a percentage of total power for a given number of memory accesses per second. This figure demonstrates that for applications that rarely access the memory, leakage energy can dominate the memory’s energy consumption, and the additional leakage energy due to the scratchpad memory can be more significant than its access energy savings.

### 3.4.3 Access Energy and Memory Size

The benefit of the scratchpad memory, in terms of active energy reduction, is determined by the difference between *SFE* and *MFE*. Therefore, a key assumption of scratchpad memories is that small memories have lower access energies than large memories. To verify this assumption, access energy data were collected from CACTI cache simulations: an older version of CACTI [61] for 180 nm and CACTI 5.3 [59] for 90, 65, 45 and 32 nm. CACTI 5.3 uses technology information projected from the 2005 ITRS [59,60]. CACTI projections are based on a real physical layout, which includes detailed device and wire models. Even if device characteristics change over time, if the layout structures are similar, the trend of smaller memories consuming less power should be consistent. The data from CACTI 5.3 are more reliable than from the older version of CACTI, because its models are more closely derived from real physical structures. The normalized access energies of 64 B to 4 kB low standby power memories for these processes are shown in Figure 3.2. The simulated measurements show that access energy generally decreases with memory size in all of the



**Figure 3.1** Leakage power for different memory access rates. Leakage power is divided by total memory power. Data are generated using CACTI (a cache and memory access time model) 5.3 for 65 nm memories optimized for low standby current [59].

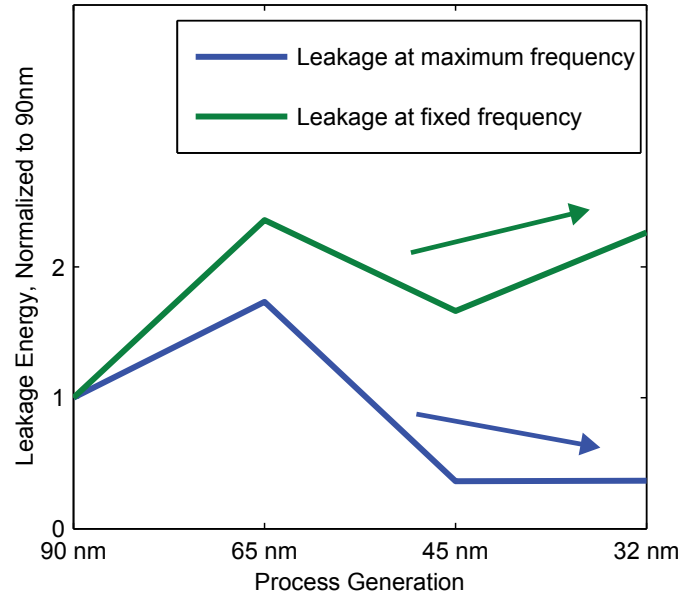


**Figure 3.2** Access energies of memories in different process technologies. Data are for low standby power memories generated using CACTI and are based on ITRS (the International Technology Roadmap for Semiconductors) 2005 projections.

simulated processes. There are a few exceptions to this rule, because CACTI simulates real layouts for memories that are primarily optimized for area and latency, rather than access energy. However, the general rule that access energy decreases with memory size is expected to hold, because larger memories will usually have longer wires and higher switching capacitance than small memories, so access energy is expected to decrease with memory size in all processes.

#### 3.4.4 Leakage Energy and Process Scaling

An increase in total leakage current is the primary drawback of a scratchpad memory, so it is important to evaluate how scaling affects leakage energy to predict how beneficial scratchpad memories will be in future processes. The evaluation of leakage energy scaling trends depends on how often a memory will be accessed for a particular application. The higher frequencies enabled by newer processes help to decrease leakage energy per clock cycle, because the clock cycles are shorter, but if an application only requires a fixed number of memory accesses per second (lower than the maximum number of accesses), then the maximum speed of the memory is irrelevant. An application that accesses memory fewer times per second than the maximum possible frequency might be equivalently viewed as having a fixed memory access frequency lower than the maximum. The leakage trend for such an application running at a frequency fixed across process generations (and below the maximum of any of those processes) is contrasted to a memory running at maximum frequency of each process in Figure 3.3. The trend for an application with a frequency that



**Figure 3.3** Leakage trends for maximum and fixed access frequencies. Leakage energy is per access cycle for different processes with the maximum number of memory accesses for the process (blue) and a fixed number of memory accesses per second lower than the maximum frequency of any of the examined processes (green). Arrows indicate the general scaling trend of each type of application.

is fixed across process generations does not improve as much with process scaling as an application that takes advantage of the maximum process frequency.

The purpose of Figure 3.3 is only to demonstrate the concept that leakage trends for fixed frequency applications are worse than trends for applications running at maximum frequencies if the newer processes target higher frequencies. Figure 3.3 is not intended to project specific leakage trends, because the 65 nm and later nodes had not yet been released when the source of its data, the 2005 ITRS, was published. A better description of recent leakage trends will be made in Section 3.5 using more recently published foundry process reports on transistor characteristics. The CACTI simulator used to generate Figure 3.3 has not yet been updated using the newest transistor data, so leakage trends must be projected using transistor data instead of simulations of memories.

### 3.5 Scaling Trends

Processes developed specifically for SoC applications have lower-power transistors and a larger variety of process options. Because embedded systems that are likely to benefit from a scratchpad memory are also likely to be built into these processes, these are the processes whose transistor characteristics should be examined.

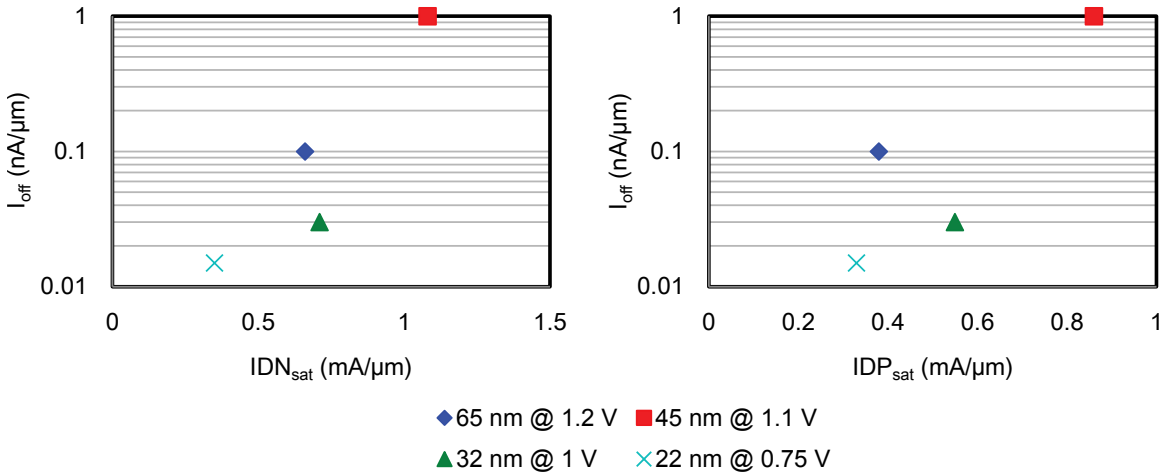
For simplicity, only Intel's processes will be presented. Similar figures could be created

for other foundries with low-power SoC processes, but the results would be similar. Figure 3.4 shows the reported achieved  $ID_{sat}$  and  $I_{off}$  characteristics for the minimum channel length of the lowest available power type of both NMOS and PMOS transistors in Intel’s 65, 45, 32 and 22 nm low-power processes [62–65]. These are low-power SoC variants of high-performance processes. The 65 nm node was Intel’s first process node to introduce an SoC process. The  $ID_{sat}$  parameter reflects the drive strength of the transistor: larger  $I_{on}$  corresponds to a faster, higher-performance transistor in saturation; lower  $I_{off}$  corresponds to a less leaky, more power-efficient transistor. The standard input voltages have decreased over time to decrease power consumption.

The overall trends illustrated in Figure 3.4 show that leakage has decreased over time. This decrease can be attributed to many process innovations, which will be examined in detail.

### 3.5.1 Process Innovations

Constant field scaling [66] has not been followed faithfully, but over recent process generations, voltages have been reduced. The last several process generations have employed a variety of innovative techniques to improve performance, while reducing leakage. An examination of the variety of process innovations that have driven improvement over the last several process generations indicates that future processes may also improve performance and reduce leakage through further innovations.



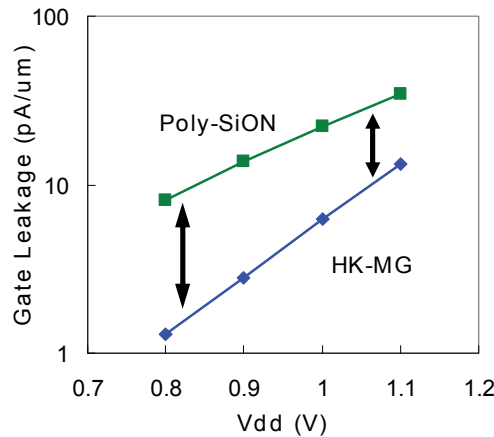
**Figure 3.4** Drive current and leakage across processes. NMOS (left) and PMOS (right) reported drive currents and off-state leakages for minimum-length transistors of the lowest available power type available in each of Intel’s 65, 45 and 32 nm SoC processes [62–65].  $ID_{sat}$  is the saturation drain current, or the current a transistor can provide when it is fully on.  $I_{off}$  is the leakage current that is present when a transistor is turned off.

Contrary to constant field scaling rules, Intel's 65 nm SoC process increased gate oxide thickness to decrease gate leakage. The good doping profile and source/drain spacers were optimized to prevent an increase in source-drain leakage, but this also increased the threshold voltage and decreased performance. The performance decrease was mitigated by uniaxial strained silicon [62].

The 45 nm SoC process introduced a dramatic structural change: a Hafnium-based high-K metal gate (HKMG), which increased performance and decreased gate leakage by allowing for a physically thicker gate with the equivalent capacitance of a thin gate. Figure 3.5 shows how dramatically an HKMG can reduce gate leakage. However, gate leakage is still much less than source-drain leakage, and Figure 3.4 indicates that although the HKMG was introduced at the 45 nm node; total  $I_{off}$  dropped much more significantly at the 32 nm node. The 32 nm node also utilized the HKMG and was the first Intel process node to use immersion lithography for the critical layer, allowing for resolution enhancement of that layer [63, 64].

Intel's 22 nm process introduced a significant advancement in process technology: a tri-gate transistor, also known as a FinFET. These transistors have a much steeper subthreshold slope than the planar 32 nm transistors [65]. This steeper subthreshold slope could be used to either significantly reduce leakage energy or to target a lower threshold voltage to increase performance. As with strained silicon and HKMG innovations, as the tri-gate transistor structure is refined, its impact will increase over time.

Research indicates that these transistor improvements are carrying over to SRAM improvements. For example, in the 22 nm technology, Intel created a low standby power



**Figure 3.5** Effect of high-K metal gate on gate leakage. The addition of a high-K metal gate decreases leakage across all voltages in an IBM 32 nm process. Image taken with permission from [67] ©2008 IEEE.

SRAM with only 10-pA/cell standby leakages [65].

Other foundries have proposed other innovations for transistor structures that are targeted at reducing leakage energy, such as partially or fully-depleted SOI (silicon on insulator) or SuVolta's deeply depleted channel (DDC) transistor [68]. In new process technologies, the improvements to energy savings and performance are now as much a result of structural redesign and innovation as they are of device dimension and voltage scaling.

### 3.5.2 SRAM Cell Design

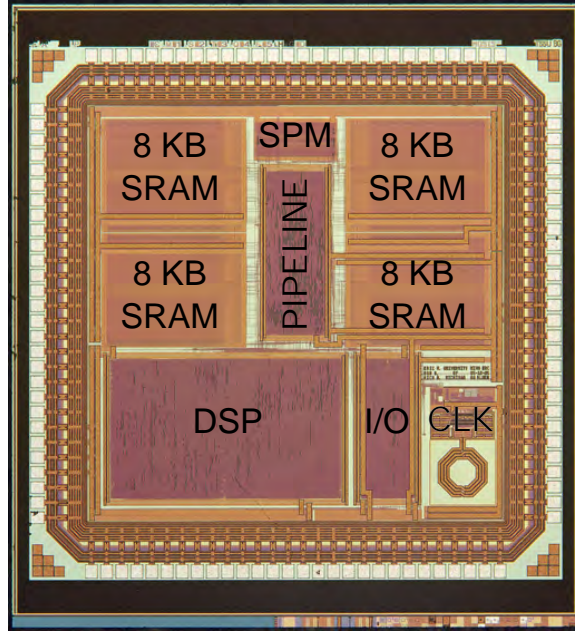
Even if the process trends do not continue to reduce leakage power, as predicted in the previous section, SRAM cell design can be modified to reduce leakage. The traditional six-transistor (6T) bit cell suffers from low noise margins, so 8T-, 9T- and 10T-bit cells have been proposed to increase noise margins and enable lower supply voltages than are possible with the 6T cell. Because 8T- or 9T-bit cells can operate at lower voltages, the increased leakage energy due to higher transistor counts is offset by the decrease in leakage per transistor. Increased noise margins can enable higher threshold transistors to be used to reduce leakage [69]. Very low leakage levels have been achieved by 9T and 10T cells at very low voltages [70, 71]. The potential for implementing these design innovations further increases our confidence that leakage energy will not come to dominate power consumption to such an extent that the increased leakage of a scratchpad memory overwhelms its dynamic power savings for the foreseeable future.

## 3.6 WIMS Microcontroller

The WIMS microcontroller has been implemented in both 180 and 65 nm processes [39, 72], as depicted in Figures 3.6 and 2.1, with a small scratchpad memory to augment the on-chip main memory banks. Because it has had relatively few architectural changes between generations, it is a good test bed for evaluating the advantages of scratchpad memories as processes are scaled. Measured data from both implementations will be used to investigate the implications of process differences on the energy savings available from small scratchpad memories, and Intel's transistor data will be used to project energy savings for a hypothetical 32 nm version of the microcontroller.

The WIMS microcontroller is designed for low-power embedded applications, such as environment and health monitoring. Its scratchpad memory is 1 kB. This memory is entirely software-controlled, but is managed in much the same way as a loop cache by the WIMS compiler. The microcontroller has a three-stage pipeline and runs at more than





**Figure 3.6** A die photo of the 180nm WIMS microcontroller.

200 MHz; the maximum frequency is unknown due to tester limitations. Its on-chip main memory is 32 kB, composed of four 8 kB banks.

Although the scratchpad memory in the WIMS microcontroller was originally referred to as a loop cache [41], it will be referred to as a scratchpad memory here to be consistent with conventional terminology [53], because it does not have a hardware controller or tag comparison circuitry. Unlike a typical cache or a hardware-controlled loop cache, the scratchpad is memory-mapped, so that it can be managed by the compiler.

### 3.6.1 Measured Data

Measurements from the 65 nm WIMS microcontroller were recorded at the University of Utah using the Verigy 93000 PS400 SOC tester. The memories are in a separate power domain, so that their power consumption could be easily separated from that of the core. To estimate energy savings, three variables were measured: main memory fetch energy ( $MFE$ ), scratchpad fetch energy ( $SFE$ ) and leakage energy ( $LE$ ). Scratchpad leakage energy ( $SLE$ ) was estimated by using memory compiler reports for a commercial 65 nm memory compiler. The memory compiler reports estimate the leakage power of the memory and showed that  $SLE$  is approximately 3% of  $LE$ , while the main memory leakage energy ( $MLE$ ) is 97% of  $LE$ . This result is intuitive, because the scratchpad memory is approximately 3% of the system memory. To measure  $MFE$  and  $SFE$ , a loop of code was executed from main memory and subsequently from scratchpad memory, and memory power was measured

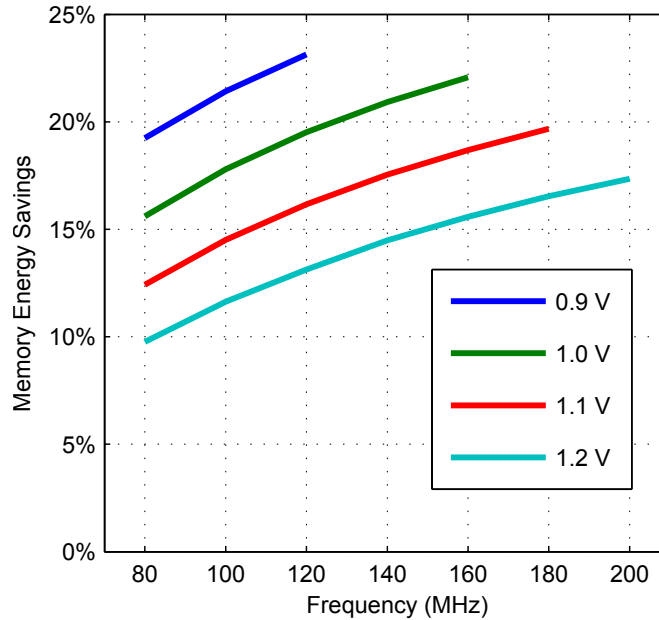
while running at various frequencies and voltages. The static memory power draw was also measured for each voltage. The power draw was divided by the clock cycle to find the energy consumed per clock cycle.  $MFE$  and  $SFE$  were found by subtracting  $LE$  from the energy measured during execution from the main memory and scratchpad memory, respectively.

Energy savings were then calculated using (3.2), but simplified by setting  $AR$  to one, because the entire program was placed in the scratchpad memory:

$$1 - ES = \frac{SFE + MLE + SLE}{MFE + MLE} \quad (3.4)$$

The percentage of energy saved using the scratchpad memory *versus* main memory in the 65 nm WIMS microcontroller implementation is depicted in Figure 3.7 for voltages between 0.9 and 1.2 V and frequencies between 80 and 200 MHz. As explained above, the total energy includes leakage energy. The scratchpad memory saves 10%–25% of total energy across all measurements, with the maximum savings achieved at the lowest voltages. These large total energy savings justify the use of a scratchpad memory in any application that has frequently used code or data and where dynamic power is a major contributor to total power.

Even in efficient operating regimes, using the lowest possible supply voltage for a given clock frequency, leakage is responsible for approximately 45% of the total energy for the



**Figure 3.7** Measured memory energy savings from executing code from the scratchpad memory instead of the main memory on the 65 nm WIMS microcontroller.

65 nm WIMS microcontroller. This relatively high leakage ratio limits the total energy savings, because the scratchpad memory can only reduce dynamic energy consumption.

### 3.6.2 32 nm Projections

To investigate whether recent scaling trends are advantageous or detrimental to scratchpad memories, the energy savings of a hypothetical 32 nm version of the WIMS microcontroller are projected. The 32 nm process was selected instead of the 22 nm process, because the 32 nm process is planar, so the assumptions necessary for the projection are more justifiable. To project the values of  $SFE$ ,  $MFE$ ,  $SLE$  and  $MLE$  that would likely be attained by a 32 nm version of the WIMS microcontroller, the energy measurements from the 65 nm version are scaled according to the transistor trends depicted in Figure 3.4. First, transistor parameters must be related to the energy variables in (3.4). Leakage energy per clock period ( $LE$ ) is proportional to the supply voltage ( $V$ ) multiplied by the leakage current ( $I_{leak}$ ) and the clock period ( $t_f$ ):

$$LE \propto V \cdot I_{leak} \cdot t_f \quad (3.5)$$

Leakage current can be further broken down into leakage per micron of gate width ( $I_{leak}/\mu\text{m}$ ) multiplied by the average gate width ( $G_W$ ):

$$LE \propto V \cdot \frac{I_{leak}}{\mu\text{m}} \cdot G_W \cdot t_f \quad (3.6)$$

Because the scratchpad memory primarily saves active energy, at the expense of a small increase in leakage energy, projected results are conservative if the decrease in leakage is underestimated and the decrease in active energy is overestimated. Based on the data presented in Figure 3.4,  $I_{leak}/\mu\text{m}$  in 32 nm is 30% of its value in the 65 nm process, while the voltage is reduced from 1.2 V to 1.0 V. To keep results conservative, no reduction in clock period was assumed. Therefore, in a 32 nm process,  $LE/G_W$ , relative to the 65 nm process, is approximately:

$$\begin{aligned} \frac{LE_{32}/G_{W32}}{LE_{65}/G_{W65}} &= \frac{I_{leak32}/\mu\text{m}}{I_{leak65}/\mu\text{m}} \cdot \frac{V_{32}}{V_{65}} \\ &= 0.3 \cdot \frac{1.0 \text{ V}}{1.2 \text{ V}} = 0.25 \end{aligned} \quad (3.7)$$

where the “32” and “65” subscripts indicate whether the parameters are for 32 nm or 65 nm, respectively. To estimate  $G_{W32}/G_{W65}$ , it is assumed that the reduction in gate area ( $A_{32}/A_{65}$ ) is approximately the same as the reduction in bit cell area. For the lowest power option, minimum  $G_L$  is 55 nm in Intel’s 65 nm technology and 46 nm in Intel’s 32 nm

technology. The bit cell area decreased from  $0.68 \mu\text{m}^2$  in 65 nm to  $0.171 \mu\text{m}^2$  in 32 nm (to 25% of its original value) [62, 64]. Therefore,

$$\begin{aligned} \frac{G_{W32}}{G_{W65}} &= \frac{A_{32}}{A_{65}} \cdot \frac{G_{L65}}{G_{L32}} \\ &= \frac{0.171 \mu\text{m}^2}{0.68 \mu\text{m}^2} \cdot \frac{55 \text{ nm}}{46 \text{ nm}} = 0.3 \end{aligned} \quad (3.8)$$

In sum,  $LE$  in 32 nm, relative to 65 nm, is estimated to be:

$$\begin{aligned} \frac{LE_{32}}{LE_{65}} &= \frac{LE_{32}/G_{W32}}{LE_{65}/G_{W65}} \cdot \frac{G_{W32}}{G_{W65}} \\ &= 0.25 \cdot 0.3 = 0.075 \end{aligned} \quad (3.9)$$

Now that the leakage scaling ratio has been found, the fetch energy scaling ratio will be considered. The fetch energy ( $FE$ ) is assumed to be proportional to dynamic switching energy, which is proportional to the switching capacitance of the wires and transistors ( $C$ ) times the supply voltage ( $V$ ) squared:

$$FE \propto C \cdot V^2 = (C_t + C_w) \cdot V^2 \quad (3.10)$$

where  $C_t$  represents capacitance due to transistors and  $C_w$  represents the capacitance due to wires. First, the division between wire capacitance and transistor capacitance should be estimated for our SRAM. This division was studied for wires of different lengths in a 130 nm process [73]. This study can be used to indicate the division between  $C_w$  and  $C_t$  if the WIMS microcontroller SRAM's wire lengths can be estimated and then converted to the 130 nm equivalent. The 65 nm WIMS SRAM uses 8 kB banks, each of which are  $320 \mu\text{m} \times 170 \mu\text{m}$ . Its wires are therefore up to  $320 \mu\text{m}$ , and these lengths should be doubled to up to  $640 \mu\text{m}$  to be equivalent to a 130 nm process. Reference [73] indicates that for wires up to  $640 \mu\text{m}$  long,  $C_w$  is approximately 40% of total  $C$ .

$C_t$  can be estimated by assuming that  $C_t$  is roughly proportional to the gate area of the transistor and that the decrease in the gate area of the transistor is proportional to the decrease in the area of one SRAM bit cell. This last assumption was also made by (3.8). The decrease in the SRAM bit cell area is from  $0.68 \mu\text{m}^2$  in 65 nm to  $0.171 \mu\text{m}^2$  in 32 nm. Therefore,

$$\frac{C_{t32}}{C_{t65}} = \frac{0.171 \mu\text{m}^2}{0.68 \mu\text{m}^2} = 0.25 \quad (3.11)$$

To estimate  $C_{w32}/C_{w65}$ , the predictive technology model was used to find the wire capacitance in the 65 nm process [74], and 2011 ITRS parameters were used to estimate the wire capacitance in the 32 nm process [75]. Using the typical wire dimensions suggested

by the predictive technology model, a typical wire in 65 nm has about 1.46 pF/cm of capacitance. Table INTC2 in the Interconnect section of the 2011 ITRS lists intermediate wires having 1.8 pF/cm of capacitance in 2011, the year the 32 nm node was introduced. This is a slight increase, but the length of each wire is halved, so the total change in  $C_w$  is:

$$\frac{C_{w32}}{C_{w65}} = 0.5 \cdot \frac{1.8 \frac{\text{pF}}{\text{cm}}}{1.46 \frac{\text{pF}}{\text{cm}}} = 0.62 \quad (3.12)$$

Although the scaling of the wire capacitance is worse than transistor capacitance scaling, this poor scaling actually makes scratchpad memories more attractive, because it increases dynamic power (which scratchpad memories reduce) relative to leakage power. Now, it is possible to find the overall capacitance scaling:

$$\begin{aligned} \frac{C_{32}}{C_{65}} &= \frac{C_{t32} + C_{w32}}{C_{65}} \\ &= \frac{C_{t32}}{C_{65}} + \frac{C_{w32}}{C_{65}} \\ &= \frac{C_{t65}}{C_{65}} \cdot \frac{C_{t32}}{C_{t65}} + \frac{C_{w65}}{C_{65}} \cdot \frac{C_{w32}}{C_{w65}} \\ &= 0.6 \cdot 0.25 + 0.4 \cdot 0.62 = 0.398 \end{aligned} \quad (3.13)$$

$FE_{32}/FE_{65}$  can be calculated using (3.10):

$$\begin{aligned} \frac{FE_{32}}{FE_{65}} &= \frac{C_{32}}{C_{65}} \cdot \left( \frac{V_{32}}{V_{65}} \right)^2 \\ &= 0.398 \cdot \left( \frac{1.0 \text{ V}}{1.2 \text{ V}} \right)^2 = 0.276 \end{aligned} \quad (3.14)$$

Using these scaling figures for  $LE$  and  $FE$ , the energy savings of the WIMS scratchpad memory will be projected for 32 nm, based on the 65 nm measured data.

### 3.6.3 WIMS Scratchpad Memory Approach

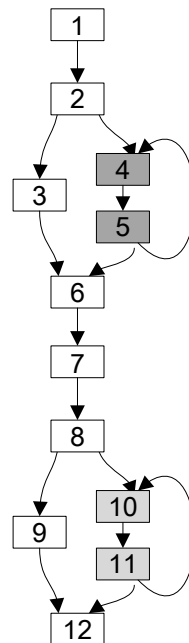
The WIMS scratchpad memory is compiler-controlled. When use of the scratchpad memory was analyzed [41], two compiler management methods were implemented: “static” and “dynamic.” The static method identifies the most frequently used loops and data objects and places as many as fit into the scratchpad memory for the duration of the program execution. The dynamic method uses a trace-selection algorithm to determine whether the energy savings of having a certain program loop in the scratchpad memory is worth the energy of copying it into the scratchpad. Copy instructions are then inserted in

the program to ensure that the selected loops or data are copied to the scratchpad memory before they are executed or accessed.

The example program depicted in Figure 3.8 will be used to illustrate the differences between the static and dynamic methods, after each algorithm is summarized. The compiler routine for the static allocation algorithm can be summarized as follows:

1. Identify the most executed blocks of code.
2. Assign these blocks of code, in the order of the most used first, to the scratchpad memory, until it is full.
3. Redirect instruction fetch (when entering code block), subroutine calls and references to these blocks to the new locations in the scratchpad memory.

In the example program in Figure 3.8, the loop consisting of Blocks 4 and 5 is the most frequently executed code in the program, followed by the loop consisting of Blocks 10 and 11. If only two blocks can fit in the scratchpad memory, then the static method will place Blocks 4 and 5 in the scratchpad memory for the duration of the execution of the program.



**Figure 3.8** Example program for explaining scratchpad management. Each numbered block represents a piece of code and an arrow represents a possible execution path. The boxes shaded the darkest are the most frequently executed. The static allocation method will hold Blocks 4 and 5 in the scratchpad memory for the duration of the execution of the program, whereas the dynamic allocation method will replace those blocks with Blocks 10 and 11 prior to their execution.

In the dynamic method, the compiler uses profiling to determine how much energy would be saved by moving a block to the scratchpad memory. If the compiler determines that the energy saved by running a block from the scratchpad memory is worth the energy of copying the code to the scratchpad memory, then it will assign those code blocks to addresses in the scratchpad memory. The dynamic allocation algorithm can be summarized as:

1. Identify blocks of code, such that the energy to copy the code to the scratchpad memory is less than the energy saved by executing the code from the scratchpad memory.
2. Assign these blocks of code addresses in the scratchpad memory and redirect references to them to their new addresses.
3. Insert copy instructions immediately prior to each of these blocks, so that the block will be in the scratchpad memory prior to its execution or prior to its being called if it is a subroutine.
4. Consolidate copy instructions to avoid unnecessary copying.

In the example program in Figure 3.8, if the compiler determines that the energy savings of running Blocks 4, 5, 10 and 11 is worth copying them to the scratchpad memory, then each block will be placed in the scratchpad memory and instructions will be inserted to copy each block to the scratchpad memory prior to its execution. Assuming two blocks can fit in the scratchpad memory, the copy instructions for Blocks 4 and 5 will be consolidated, as well as the copy instructions for Blocks 10 and 11.

The compiler uses execution profile information and an energy benefit heuristic to determine which blocks to move to the scratchpad memory. Details of the compilation procedures used in the static and dynamic allocation methods are beyond the scope of this paper, but are available in [41].

The energy savings of the dynamic and static methods will be presented for the 180 and 65 nm processes based on measured data, and the energy savings of a hypothetical 32 nm process will be projected.

#### 3.6.4 Static Benchmarks

The benchmarks used to compare the scratchpad memory energy savings in the 180, 65 and 32 nm processes are the same as those used to compare the static and dynamic allocation methods in [41]. Those benchmarks are a subset of the MediaBench [76] and MiBench [77]

benchmarks, oriented towards embedded systems. This selection of benchmarks is convenient, because scratchpad memory access rates for each of these benchmarks have already been published in [41] for the WIMS microcontroller, using the static allocation method. These data can be used to project the energy savings ( $ES$ ) of the microcontroller using the model represented by (3.2), repeated here with a slight modification for convenience:

$$1 - ES = \frac{AR \cdot SFE + (1 - AR) \cdot MFE + LE}{MFE + MLE} \quad (3.15)$$

Leakage energy ( $LE$ ) has replaced  $MLE + SLE$  for brevity ( $LE = MLE + SLE$ ). The access rates ( $AR$ ) are found in [41].  $SFE$ ,  $MFE$  and  $LE$  were measured for 0.9 V and 120 MHz, as described in Section 3.6.1, and  $MLE = 0.97 \cdot LE$ , as noted in the same section. The frequency of 120 MHz was selected, because it was the highest operational frequency for the microcontroller with a 0.9-V supply voltage. Projected energy savings are presented in Figure 3.9(a) for a 256 B scratchpad memory. The 65 nm WIMS microcontroller has a larger 1 kB scratchpad memory, but scratchpad access rates were not available for all benchmarks for that larger size, so the results presented are conservative. The 32 nm projections of each of these variables were made using the method described in Section 3.6.2.

The scratchpad memory is limited to saving dynamic power, which is only  $\sim 55\%$  of the WIMS microcontroller's total memory power. The remaining  $\sim 45\%$  is due to leakage. Thus, despite savings as high as 40% in terms of dynamic energy consumption, Figure 3.9(a) shows lower total energy savings because of the high leakage power.

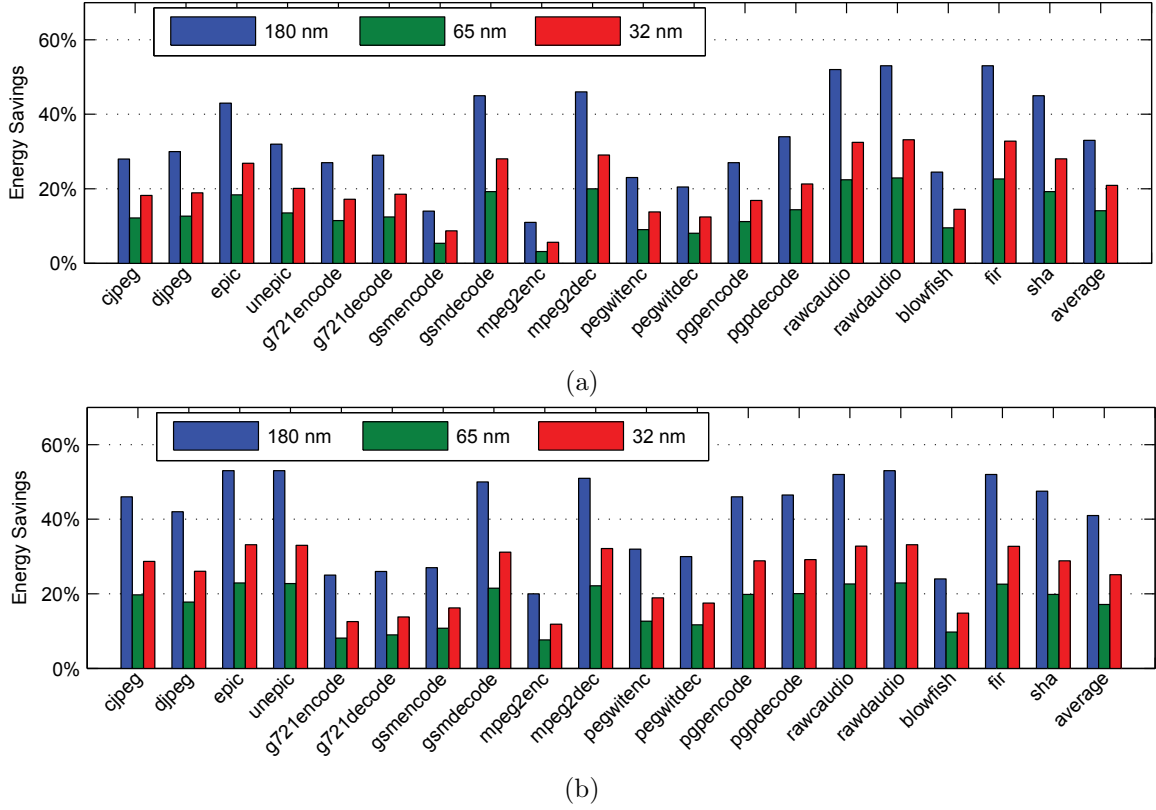
### 3.6.5 Dynamic Benchmarks

To use the measured data to project energy savings from using the dynamic allocation scheme, two new variables are added to the model:

$$1 - ES = \frac{AR \cdot SFE + (1 - AR) \cdot MFE + CO \cdot (MFE + SWE) + LE}{MFE + MLE} \quad (3.16)$$

$CO$  (copy overhead ratio) represents the impact of the additional instructions necessary to copy code from the main memory to the scratchpad memory during execution, per memory access. Because a copy involves a read from main memory and a write to scratchpad memory,  $CO$  is proportional to  $MFE$  and  $SWE$  (scratchpad write energy). For the 180 nm processor, all of these variables are reported for each benchmark except for  $CO$ , which was solved for using (3.16). This variable was usually very small, often  $<1\%$ , with a maximum value (for the g721encode benchmark) of 7%. Its impact was overwhelmed by the increase in  $AR$  which occurs when using the dynamic allocation scheme.





**Figure 3.9** Comparison of energy savings from a scratchpad memory in 180, 65 and 32 nm technologies, using (a) the static allocation method and (b) the dynamic allocation method.

For the 65 nm processor and 32 nm projection, the copy overhead ratio  $CO$  is the same as it is in the 180 nm processor, because the memory sizes and ISA (instruction set architecture) are the same, so the number of copy instructions executed to fill the scratchpad during execution will be the same.  $SFE$  was substituted for  $SWE$ , because the available  $SFE$  data were more reliable. Tests on the 180 nm processor indicated that the read and write energies are similar; specifically, the read and write energies of a 512 B memory were virtually identical, and the read and write energies of a 1 kB memory (the size of the scratchpad on the WIMS microcontroller) were within 20% of each other.  $ES$  was then calculated for the 65 nm processor using (3.16). The 32 nm projections were again made using the method described in Section 3.6.2.

The results depicted in Figures 3.9 reveal that: (1) the dynamic allocation method generally increases the energy savings achieved by the scratchpad memory; (2) the scratchpad memory saves substantial energy in every examined process; and (3) that energy savings expected in a 32 nm version of the microcontroller should be even greater than observed in

the 65 nm version.

### 3.7 Conclusions

Despite the high leakage ratio at the 65 nm process node, this study has shown that the energy saved by using scratchpad memories continues to merit their consideration. In nodes with lower leakage ratios, such as the 45 or 32 nm nodes, or with memories that are optimized to minimize leakage power, scratchpad memories are even more attractive. In particular, the largest gains were found for frequent memory accesses and minimal memory supply voltage. Although this study focused on using scratchpad memories for instruction storage, we believe these conclusions will hold if the scratchpad memories are used for data storage, as well. Applications with low memory-access frequency or that use processes with high leakage-to-dynamic power ratios will see limited savings from scratchpad memories. These conclusions were supported by data measured from the WIMS microcontroller implemented in both 180 and 65 nm processes. CACTI simulations were used to demonstrate how process characteristics impact SRAM energy consumption and scratchpad memory energy savings. Intel process technology reports indicate that process innovations will continue to decrease leakage energy. These data can guide the design process at multiple levels from software to architecture to process selection. In many applications, scratchpad memories will provide significant savings in dynamic power consumption at little cost in leakage power, even in advanced semiconductor processes. This power savings is likely to continue for the process nodes of the foreseeable future.

## CHAPTER 4

### A LOW-POWER 3-STAGE 1.8 V OPERATIONAL AMPLIFIER

#### 4.1 Introduction

The opamp (operational amplifier) is an important analog sub-block that is used in many of the microsystem's analog components. Instead of designing several opamps, each optimized for a particular use in the microsystem, considerable design time can be saved by using a single opamp that has been designed to meet the requirements of many blocks. This chapter will consider the diverse uses of the opamp, its requirements, and the motivation for the selected architecture.

#### 4.2 Desired Specifications

The opamp will be the primary component or a subcomponent in many of the Utah Microcontroller's analog circuit blocks:

- Analog-to-digital converter (ADC)
- Digital-to-analog converter (DAC)
- Linear voltage regulators
- High input-impedance buffer
- Output buffer

Each of these items has different requirements, so the opamp must be overdesigned for any particular component so that it can simultaneously meet the requirements of all of these components. There will be many copies of the opamp, so to keep total system power low, the opamp must be very low power.

In the sigma-delta ADC, the opamp acts as part of the integrator in each of the ADC's two stages. The requirements for the opamp used in the integrator are derived from the

ADC requirements in Section 5.3.3.2. That section concludes that the opamp must provide  $8.8\text{ }\mu\text{A}$  slewing current but only about 40 dB gain. The ADC also requires the opamp to have rail-to-rail input and output stages so that the ADC can support rail-to-rail input voltages.

Because the DAC must drive voltages up to 1.5 V, the opamp must operate at the 1.8 V rail (instead of the 1.2 V rail). The DAC must drive large off-chip capacitances, so the opamp must be capable of driving multiple mA of current [22]. To avoid adding distortion when buffering signals, it must have high gain.

The complete opamp requirements are summarized in Table 4.1.

## 4.3 Opamp Architecture

### 4.3.1 Limitations of Classic 2-stage Opamp

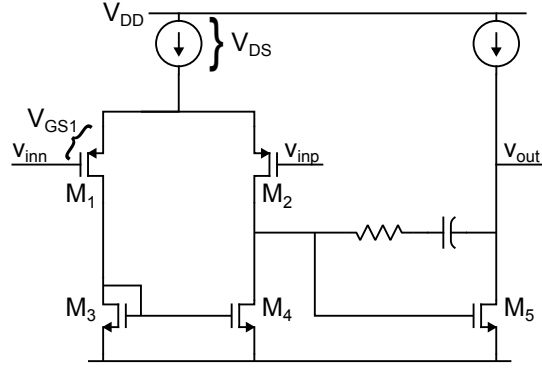
The most common design is the classic 2-stage operational amplifier, but this has some drawbacks. A simplified schematic of the 2-stage opamp is depicted in Figure 4.1. In this schematic, transistors  $M_1$  and  $M_2$  form the input differential pair. The total current flowing through the differential pair is kept constant by the current source, and the input voltages  $v_{inn}$  and  $v_{inp}$  determine how much of that current flows through each branch of the differential pair. The output currents are summed using the current mirror formed by transistors  $M_3$  and  $M_4$ , resulting in an output voltage of the first stage on the node between  $M_2$  and  $M_4$ .  $M_5$  acts as a common-source amplifier to further amplify the signal as part of the second stage.

There are drawbacks to the input stage and the output stage using this design. The input and output ranges are limited, and the drive current is also very limited. The limitation on the input voltage is demonstrated by an equation relating the input voltage to other voltages in the circuit:

$$v_{in} = V_{DD} - |V_{DS}| - |V_{GS1}| \quad (4.1)$$

**Table 4.1** Opamp requirements.

Requirement	Specification
Output Current	$>5\text{ mA}$
Input Range	Rail-to-rail (0 to 1.8 V)
Power consumption	As low as possible
Gain	$>40\text{ dB}$
Offset	$<5\text{ mV}$



**Figure 4.1** A schematic of the classical 2-stage opamp. This opamp does not meet the opamp requirements, and a consideration of its limitations demonstrates the need for different input and output stages.

$V_{DD}$  is the supply voltage,  $V_{DS}$  is the drain to source voltage of the transistor acting as the current source, and  $V_{GS}$  is the gate-to-source voltage of the input transistors,  $M_1$  and  $M_2$ . For the opamp to operate properly, all transistors must be kept in saturation. For the input transistors to be saturated,  $V_{GS}$  must be more than the threshold voltage  $V_T$ , so the equation can be rewritten as:

$$v_{in} \leq V_{DD} - |V_{DS}| - |V_T| \quad (4.2)$$

$V_T$  is a process characteristic that cannot be revealed under a nondisclosure agreement, but IBM's report of a similar 65 nm process showed that  $V_T$  for standard and high-threshold 1.2 V transistors was 0.4 V and 0.6 V, respectively [78]. Although our opamp will use the 1.8 V I/O transistors, these figures give the reader an idea of the ballpark figure for  $V_T$ . If  $V_{DS}$  is only 0.1 V,  $V_{DD}$  is 1.8 V, and  $V_T$  is only 0.5 V (selected only for an example calculation, this may not be accurate for our process), then maximum  $v_{in}$  is:

$$v_{in} \leq 1.8 \text{ V} - 0.1 \text{ V} - 0.5 \text{ V} = 1.2 \text{ V} \quad (4.3)$$

This is a major limitation of the classic 2-stage opamp. The output drive current is also limited. Consider the case in which  $v_{out}$  swings high. The current draw of the output transistor,  $M_5$ , will approach zero. The entire quiescent bias current for the output stage is then directed to the output. However, the opamp cannot provide drive current that is greater than the bias current. This means that the opamp cannot provide an output drive current that is much greater than its own typical power consumption, as is desirable for our target applications.

### 4.3.2 A Suitable Input Stage

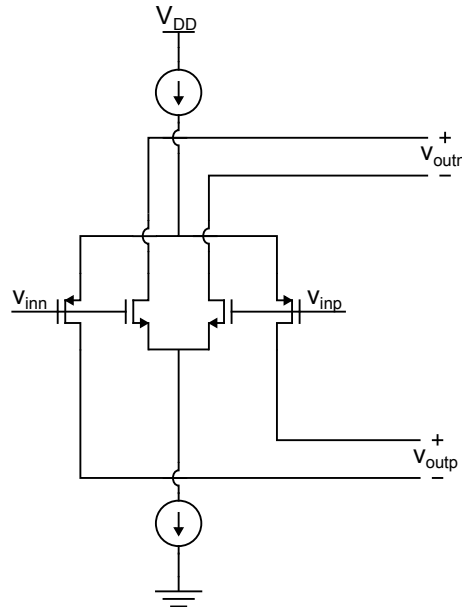
Designers have found ways to overcome these limitations. One key observation is that the limited voltage range of the 2-stage opamp does not apply to the lower voltage limit. Starting from ground in Figure 4.1 and working up through  $M_3$  or  $M_4$  and then to the gates of  $M_1$  and  $M_2$ , we find:

$$v_{in} = |V_{DS}| + |V_{DS}| - |V_{GS1}| \quad (4.4)$$

Using the previous estimates of  $V_{DS}$  and  $V_T$ , the lower limit of  $v_{in}$  is

$$v_{in} \geq 0.1 \text{ V} + 0.1 \text{ V} - 0.5 \text{ V} = -0.3 \text{ V} \quad (4.5)$$

The opamp can operate even with the input voltage below ground. If the upper input voltage limit must be at or above  $V_{DD}$ , the opamp architecture can exchange NMOS transistors for PMOS transistors, so the the input transistors are NMOS. Using these insights, designers found that a full input voltage range can be supported by placing an input NMOS pair in parallel with an input PMOS pair, as shown in Figure 4.2. In this schematic, the inputs are connected to both the NMOS and PMOS transistors. The current outputs of the two differential pairs are summed together using a summing circuit.



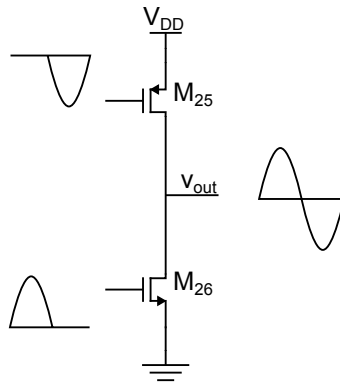
**Figure 4.2** A schematic of a full rail-to-rail input stage with NMOS and PMOS input differential pairs in parallel.

is with a class AB output stage, also referred to as a push-pull stage. A simplified version is shown in Figure 4.3. This output stage works by turning on only one transistor at a time to prevent short-circuit current through the transistors. The waveform on the left shows the input voltage. When the  $v_{in}$  swings low, the top transistor turns off, so that the output voltage is controlled by the bottom transistor, and vice versa. If the output voltage is limited through negative feedback, then the transistor will turn on just enough to drive the output to the desired voltage. Using this circuit, the output drive current can be much greater than the current typically drawn by the output stage.

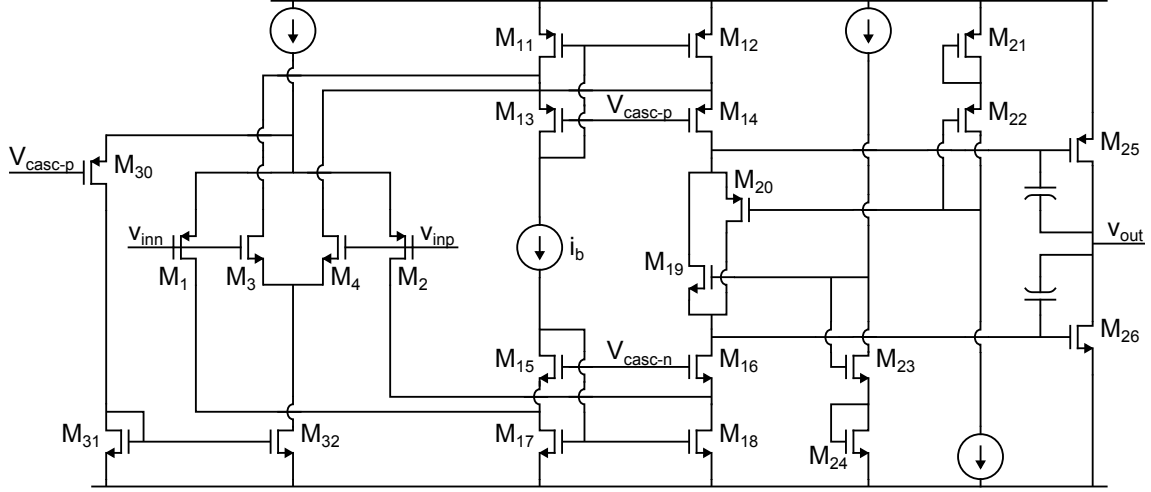
#### 4.3.4 Opamp Schematic and Operation

Many designs have been proposed in the literature, and using the above considerations as guidelines, the schematic depicted in Figure 4.4 was selected. A more detailed view using only transistors is given later in the chapter, but Figure 4.4 will be more useful for understanding the principle of operation of the opamp. This opamp architecture is described in Chapter 9 of [80] and Chapter 5 of [81].

The input stage is two differential pairs in parallel, which has the advantages described in Section 4.3.2. Three transistors have been added to the input stage:  $M_{30}$  to  $M_{32}$ . These three transistors are used to hold a relatively constant  $g_m$  (or transconductance) across the common-mode input range of the opamp. If  $v_{inn}$  and  $v_{inp}$  are both high, then  $M_1$  and  $M_2$  may both shut off, so the  $g_m$  of the opamp is reduced compared to when both sets of input transistors are in the saturation region. A similar problem occurs when the common-mode voltage (the average value of the inputs) is low enough to turn off the NMOS transistors,



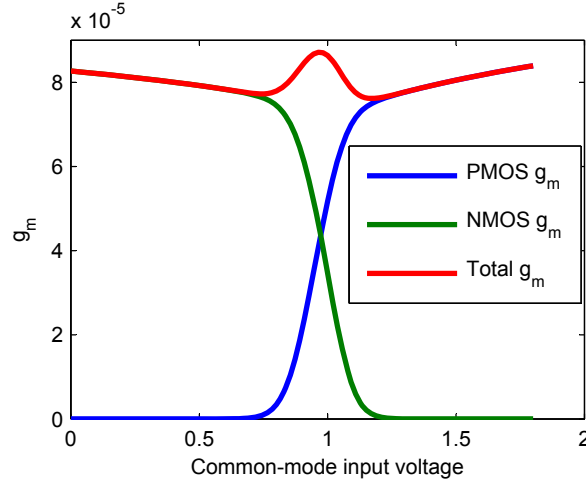
**Figure 4.3** Simplified schematic of class AB output stage. The left waveform shows the input voltage, the top and bottom show the current through the top and bottom transistors, respectively, and the right waveform shows the output current or voltage. Figure derived with permission under Creative Commons license from [79].



**Figure 4.4** Schematic of low-power opamp.

$M_3$  and  $M_4$ . To correct this problem,  $M_{30}$  is used to capture whatever current is not drawn by  $M_1$  and  $M_2$ , and a current mirror of  $M_{31}$  and  $M_{32}$  increase the current flowing through  $M_3$  and  $M_4$  by a ratio to that current. If the common-mode voltage is low enough that no current flows through  $M_{30}$ , then  $M_3$  and  $M_4$  are effectively turned off. Conversely, when the common-mode voltage is high enough that the current flowing through  $M_1$  and  $M_2$  is near zero, then the current drawn by  $M_{32}$  through  $M_3$  and  $M_4$  is at a maximum.

This arrangement results in a relatively constant  $g_m$ , as shown in simulation in Figure 4.5. The total  $g_m$  varies by less than 15% across the entire rail-to-rail input voltage



**Figure 4.5** Transconductance of input differential pairs across common-mode input voltage. The total transconductance ( $g_m$ ) of the input differential pairs is kept relatively constant across the entire common-mode input range by transistors  $M_{30}$ – $M_{32}$ .



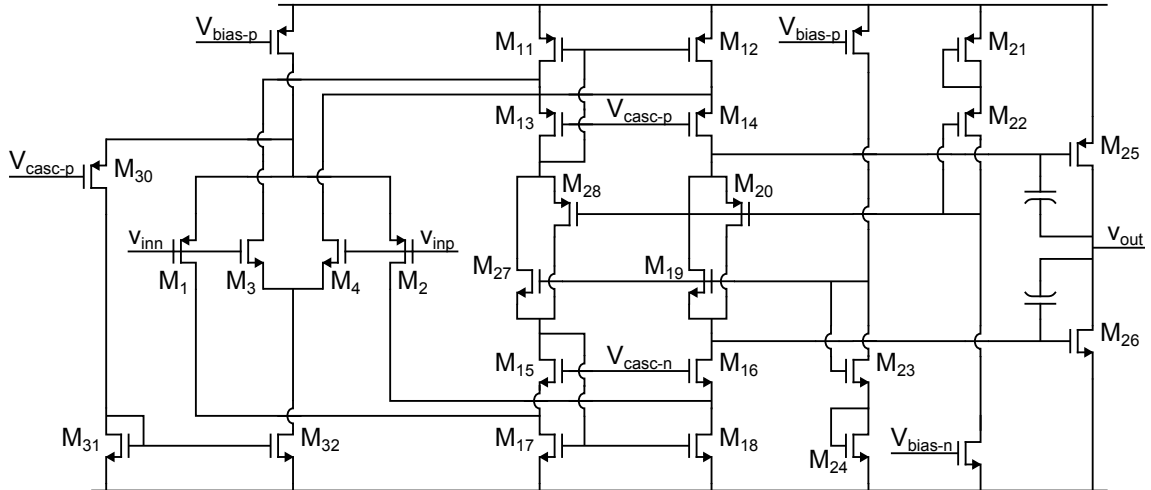
range.

Transistors  $M_{11}$ – $M_{18}$  comprise a summing circuit. The drains of  $M_1$ – $M_4$  all output a modulated current. The operation of the PMOS transistors of the summing circuit are identical to the NMOS transistors in the summing circuit, so discussion of the summing circuit will first focus on the NMOS transistors.  $M_{17}$  is diode-connected (with  $M_{15}$  as a cascode) that mirrors  $i_b$  added to the current output of  $M_1$ , which will be referred to as  $i_{M1}$  for short.  $M_{18}$  will therefore draw that same current,  $i_b + i_{M1}$ , but because some of its current comes through  $M_2$  as well as  $M_{16}$ , the current drawn through the drain of  $M_{16}$  is approximately  $i_b + i_{M1} - i_{M2}$ .  $M_{16}$  acts as a cascode to increase the output resistance of the current summing circuit.

$M_{19}$  sets the bias point of one output transistor,  $M_{26}$ , while  $M_{23}$ – $M_{24}$  set the bias point of  $M_{19}$ . The PMOS transistors function in the same way.

The full realization of the opamp is shown in Figure 4.6. The floating current source is realized with  $M_{27}$ – $M_{28}$ .

All transistors were biased to operate in the saturation region to avoid the increased variability issues with subthreshold operation [82]. The transistors, while significantly larger than minimum dimensions in a 65 nm process, were generally kept at the minimum length of a 1.8 V I/O transistor in order to minimize both dynamic and leakage power. The transistor widths were also relatively small to minimize operating current. In order to minimize the input offset, the overdrive voltage of the input transistors was kept very small [83].



**Figure 4.6** Transistor-level schematic of low-power opamp.

### 4.3.5 Noise Analysis

Fortunately only a few transistors contribute to the total noise of the opamp. After the signal has been amplified by one stage, the input-referred noise contributions of the transistors in the subsequent stages are divided by the first stage's gain, generally making their noise contribution insignificant. There are two significant sources of noise that must be considered:

- Thermal noise
- Flicker noise

The transistor thermal noise is modeled as a current source in parallel with that transistor. The noise output spectrum is:

$$\overline{I_{therm}^2} = 4kT\gamma g_m \quad (4.6)$$

according to [84]. The value of  $\gamma$  for a long-channel transistor is  $2/3$ . This power spectrum is roughly constant across the relevant frequency spectrum.

Flicker noise is modeled as a voltage source in series with the transistor gate, with a power spectrum given by:

$$\overline{v_{1/f}^2} = \frac{K}{C_{ox}WLf} \quad (4.7)$$

Conventionally the totally noise of the opamp is reported as the input-referred noise. To convert thermal noise into input-referred noise, the value of the current source in parallel with the transistor is divided by  $g_m^2$ . This results in the value of a series noise voltage source at the gate of the transistor. This can be added to the flicker noise for the total transistor noise power of a single transistor. For example, the total noise due to transistor M1 is:

$$\overline{v_{n,1}^2} = \frac{4kT\gamma}{g_{m,1}} + \frac{K_p}{C_{ox}W_1L_1f} \quad (4.8)$$

In this equation,  $g_{m,1}$  refers to the transconductance of transistor M<sub>1</sub>,  $K_p$  is the  $K$  parameter of the PMOS transistors, etc. The noise sources of each transistor are uncorrelated, so to represent the total noise as a single noise source on one of the two differential inputs, the noise sources are simply added. A more rigorous proof of this practice is given in the treatment of differential input-referred noise in [84]. Therefore, the total noise contributions of the four input transistors (M<sub>1</sub>–M<sub>4</sub>) is:

$$\overline{v_{n,1-4}^2} = \sum_{i=1}^4 \left( \frac{4kT\gamma}{g_{m,i}} + \frac{K_i}{C_{ox}W_iL_if} \right) \quad (4.9)$$

To yield insights into how the noise can be minimized from this equation,  $g_m$  should be translated into variables that are kept constant by the circuit.

$$\overline{v_{n,1-4}^2} = \sum_{i=1}^4 \left( \frac{4kT\gamma}{\sqrt{2\mu_i C_{ox} \frac{W}{L} I_d}} + \frac{K_i}{C_{ox}W_iL_if} \right) \quad (4.10)$$

Equation (4.10) yields insights into how noise can be reduced. Low frequency noise tends to be dominated by flicker noise, while at high frequencies thermal noise tends to dominate.  $K$  is a process-dependent parameter, as is  $C_{ox}$ , so the only way to decrease flicker noise indicated by (4.10) is to

**Increase  $W_i$  and  $L_i$ .** The disadvantage of that strategy is that leakage current is increased.

To decrease the thermal noise of the input transistors, there are two options:

**Increase  $I_d$ .** The drawback is increased power consumption.

**Increase  $W/L$ .** If  $W/L$  is increased too dramatically, the transistors will be biased into the subthreshold region, which increases susceptibility to process variation [82].

The input transistors are not the only transistors that contribute to the opamp noise. Fortunately, transistors M<sub>13</sub>–M<sub>16</sub> do not significantly contribute to the input-referred noise of the opamp because the gain seen at the drains of these transistors makes their input-referred noise contributions negligible. However, transistors M<sub>11</sub>–M<sub>12</sub> and M<sub>17</sub>–M<sub>18</sub> do contribute to the noise because their drains are connected directly to the drains of the input transistors. Fortunately this makes the calculation of their contribution to the input-referred noise fairly straightforward: the noise modeled by a current source at the drain should be divided by  $g_{m,i}^2$  of the input transistor each is connected to. The total input-referred noise contribution of these transistors is:

$$\overline{v_{n,11-18}^2} = \sum_{i=11,12,17,18} \left( \frac{4kT\gamma g_{m,i}}{g_{m,j}^2} + \frac{K_i \cdot g_{m,i}^2}{C_{ox}W_iL_if g_{m,j}^2} \right) \quad (4.11)$$

In this case, the subscript  $j$  refers to the input transistor that the transistor is connected to. For example, for  $i = 11$ ,  $j = 3$  because the drain of  $M_{11}$  is connected to the drain of  $M_3$ . Next the transconductances will be put in terms of the drain current.

$$\begin{aligned} \overline{v_{n,11-18}^2} &= \sum_{i=11,12,17,18} \left( \frac{4kT\gamma\sqrt{2\mu_i C_{ox} \frac{W_i}{L_i} I_{D,i}}}{2\mu_j C_{ox} \frac{W_j}{L_j} I_{D,j}} + \frac{K_i 2\mu_i C_{ox} \frac{W_i}{L_i} I_{D,i}}{C_{ox} W_i L_i f \cdot 2\mu_j C_{ox} \frac{W_j}{L_j} I_{D,j}} \right) \\ &= \sum_{i=11,12,17,18} \left( \frac{2kT\gamma\sqrt{2\mu_i \frac{W_i}{L_i} I_{D,i}}}{\mu_j \sqrt{C_{ox} \frac{W_j}{L_j} I_{D,j}}} + \frac{K_i \mu_i W_i I_{D,i}}{C_{ox} L_i^2 f \cdot \mu_j \frac{W_j}{L_j} I_{D,j}} \right) \end{aligned} \quad (4.12)$$

Unfortunately the equation is more complex in this case, but there are still some important insights as to how noise can be reduced. Flicker noise (represented by the term after the ‘+’ in the summation) can be reduced in the following ways:

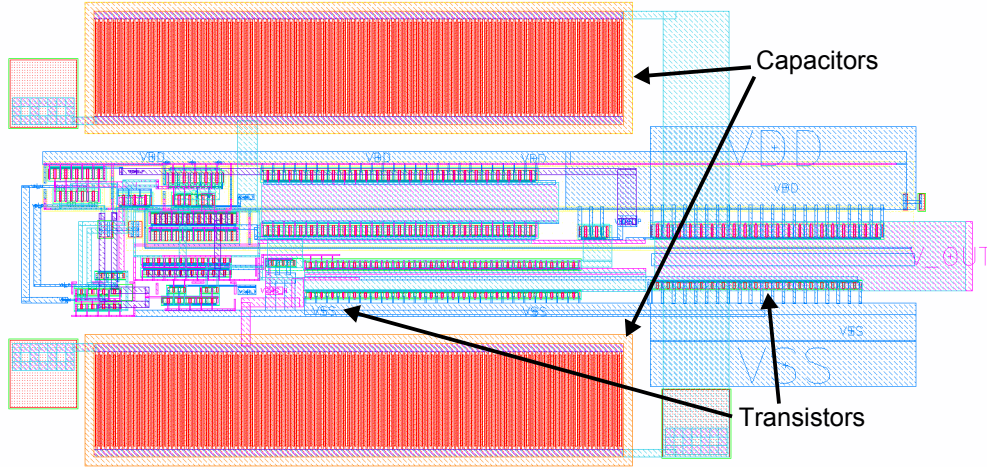
- Increase  $L_i$ . Fortunately this suggestion overlaps with the suggestion for reducing the flicker noise from the input transistor itself (increase  $W_i$  and  $L_i$ ). The drawback is increased input capacitance and therefore power consumption.
- Increase  $W_j/L_j$ . As mentioned above, if  $W_j/L_j$  is increased too dramatically, the transistors will be biased into the subthreshold region.
- Increase  $I_{D,j}$ . Any increase in bias current directly increases power consumption.

Two of these options (and their accompanying drawbacks) also apply to reducing thermal noise: increase  $W_j/L_j$  or increase  $I_{D,j}$ , as can be observed by examining (4.12).

## 4.4 Simulated and Measured Results

The layout of the fabricated opamp is shown in Figure 4.7. The simulated results, along with measured results, are listed in Table 4.2. Some metrics have been displayed previously, such as the rail-to-rail simulated transconductance shown in Figure 4.5. In simulation, the opamp achieved a complete rail-to-rail input and output. It also achieved very low power: only 81  $\mu\text{W}$ , and very high gain. These are the most significant opamp metrics. The opamp also achieved a gain-bandwidth product and phase margin that are adequate for the target applications.

It is possible to further tweak copies of the opamp for specific uses. For example, the output stage transistors could be smaller for the opamps used in the ADC integrators, to help further reduce power consumption. However, because the bias voltages of the class AB output stage described in Section 4.3.3 have been carefully optimized, the reduced power



**Figure 4.7** Layout of the fabricated opamp.

consumption may be less than anticipated. The output stage consumes only 16–45% of the power of a bias circuit and opamp, depending on the common-mode input level. Decreasing the bias current of other stages increases transistor noise, which impacts all of the target components. Simulated results showed that a single opamp design sufficiently satisfied all component requirements.

Unfortunately, the measured results did not match the simulated results very well. The fabricated opamp, and its bias circuitry, drew much less current than expected from simulation. This decreased the slew rate and gain-bandwidth, and was due to two causes:

1. A design error in the bias circuitry, in which a transistor was sized inappropriately.

**Table 4.2** Simulated and measured opamp results. Measured results include measurements without changing the biases (under “measured”), and results with the simulated bias voltages driven externally (under “biases set”).

Specification	Simulated	Measured	Biases set
Input offset	107 $\mu\text{V}$	27 mV	Low
Open loop gain	81 dB	82 dB	87 dB
Gain-bandwidth product	4.9 MHz	283 kHz	2.9 MHz
Phase margin	65°	~65°	~65°
Slew rate	.69 V/ $\mu\text{s}$	.01 V/ $\mu\text{s}$	.8 V/ $\mu\text{s}$
Supply current	45 $\mu\text{A}$	1 $\mu\text{A}$	23 $\mu\text{A}$
Drive strength	1.65 mA	1.6 mA	1.6 mA
Input referred noise		13 $\mu\text{V}/\sqrt{\text{Hz}}$ at 100 Hz 6 $\mu\text{V}/\sqrt{\text{Hz}}$ at 1 KHz	23 $\mu\text{V}/\sqrt{\text{Hz}}$ at 100 Hz 11 $\mu\text{V}/\sqrt{\text{Hz}}$ at 1 KHz

2. A layout effect which resulted in decreased effective width of the same transistor.

The bias circuit which generates the bias voltages for the opamp is depicted in Figure 4.8. It is a wide-swing constant-transconductance bias circuit, and its operation is discussed in Section 6.1 of [85]. It is based on a similar bias circuit in Section 5.2 of [85]. The author's initial understanding of this circuit was that the bias resistor sets the current coming through the left-most branch, and  $M_8$  and  $M_9$  act as a current mirror resulting in the same current flowing through  $M_7$  and  $M_6$ . The value of this current sets  $V_{GS3}$ , and results in the same amount of current flowing through each of the four branches of the bias circuit.

This initial understanding of the circuit was incomplete. It is true that the  $V_{GS}$  voltages of  $M_8$ ,  $M_9$ , and  $M_1$  match the  $V_{GS}$  voltages of  $M_7$ ,  $M_6$ , and  $M_4$ , respectively; and that  $M_8$  and  $M_9$  act as a current mirror resulting in the same current flowing through  $M_7$  and  $M_6$ ; however, the relationship between  $M_2$  and  $M_3$  is different. Applying Kirchhoff's voltage law around the loop between  $M_2$ ,  $M_3$ , and ground results in:

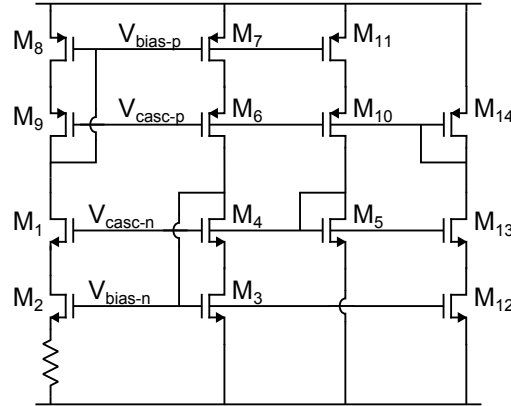
$$V_{GS3} = V_{GS2} + I_{D2}R \quad (4.13)$$

$M_2$  and  $M_3$  operate in saturation, so  $V_{eff} = V_{GS} - V_t$  and  $V_{eff} = \sqrt{\frac{2I_D}{\mu_n C_{ox}(W/L)}}$ . Applying these equations to (4.13) results in

$$\sqrt{\frac{2I_{D3}}{\mu_n C_{ox}(W/L)_3}} = \sqrt{\frac{2I_{D2}}{\mu_n C_{ox}(W/L)_2}} + I_{D2}R \quad (4.14)$$

Because  $I_{D2} = I_{D3}$ , this equation can be restated as

$$\sqrt{\frac{2I_{D3}}{\mu_n C_{ox}(W/L)_3}} = \sqrt{\frac{2I_{D3}}{\mu_n C_{ox}(W/L)_2}} + I_{D3}R \quad (4.15)$$



**Figure 4.8** Opamp bias voltage generation circuit.

which can be simplified to

$$\frac{2}{\sqrt{\mu_n C_{ox}(W/L)_3 I_{D3}}} \left[ 1 - \sqrt{\frac{(W/L)_3}{(W/L)_2}} \right] = R. \quad (4.16)$$

Using the saturation region equation  $g_{m3} = \sqrt{2\mu_n C_{ox}(W/L)_3 I_{D3}}$ , this finally results in

$$g_{m3} = \frac{2 \left[ 1 - \sqrt{\frac{(W/L)_3}{(W/L)_2}} \right]}{R} \quad (4.17)$$

Equation (4.17) shows that, to a 1<sup>st</sup>-degree approximation, the ratio of the (W/L) parameters of  $M_2$  and  $M_3$  sets the transconductance of  $M_3$ . Reference [85] recommends a ratio of 4-to-1. Unfortunately, these two transistors were sized identically. This sizing, according to (4.17), results in zero current draw. More accurate transistor simulations showed that the generated biases resulted in a supply current of 45  $\mu\text{A}$ , as listed in Table 4.2. Although the simulation models are generally very accurate, these simulations did not reveal the degree to which the bias voltages varied depending on mismatch between the widths of  $M_2$  and  $M_3$ . Also, simulations (even parasitic-extracted simulations) do not take into account layout effects on transistor performance.

One of these transistors,  $M_2$ , was laid out on the corner of the circuit, without surrounding dummy gates, and was subject to relatively severe layout effects, which reduced the effective width of that transistor. This reduced the opamp's current draw to only 1  $\mu\text{A}$ . In simulation, a 20% reduction in the width of  $M_2$  resulted in similar performance to that observed in the fabricated circuit. The transistors were also more susceptible to layout effects because they were biased to the edge of the saturation region, and close to the subthreshold region, in an attempt to minimize power consumption. The widths of the transistors were large relative to the minimum gate width of the process, but were not substantially wider than the minimum width for the thick oxide transistors necessary to support a 1.8 V power supply. These design mistakes reduced certain performance parameters.

The test circuit allowed driving in the correct bias voltages externally, and measuring the performance parameters with the correct bias voltages. With the correct bias voltages, the circuit performance was similar to simulation, allowing for some performance changes due to transistor variation, as shown in Table 4.2.

The input offset is not reported for when the biases were driven externally because slight tweaks to the bias voltages altered the input offset, such that it could be set to be very low. The fact that input-referred noise was measured to be higher when the biases

were driven externally indicates that noise was likely entering the circuit through the bias connections. If the test setup allowed for shielded connections, the measured noise would likely be reduced.

## 4.5 Conclusion

A design error and the layout effects discussed in Section 4.4 resulted in reduced opamp performance. The two performance metrics most affected were the gain-bandwidth product and slew rate. However, both measurements and simulations strongly suggest that the cause of the performance degradation has been identified and could be corrected in a revision of the opamp. Measurements listed in Table 4.2 show that with the proper bias voltages, the opamp performs close to simulation and meets all application requirements.



## CHAPTER 5

### A LOW-POWER HIGH-PRECISION SIGMA-DELTA ADC

#### 5.1 Introduction

The two target applications both require an analog-to-digital converter (ADC). The primary target application, a Smart Intravaginal Ring (S-IVR), is needed to help study the efficacy of the microbicide released from the ring in preventing the transmission of AIDS. A microsystem that could monitor the use of an IVR must have chemical sensors capable of reporting on its use, a microcontroller to read from those sensors, and an ADC to interface between the microcontroller and chemical sensors.

The S-IVR requires temperature, conductivity, and pH sensors. The temperature and conductivity sensors can indicate whether the S-IVR is in the human body, and the pH sensor indicates whether semen are present. The pH sensor directly converts the ambient fluid's pH value into an analog voltage, while the conductivity sensor allows the ambient fluid's conductivity to be measured as a resistance. An AC current source can convert the resistance of the conductivity sensor into an n AC voltage. The temperature sensor presents a resistance that can be converted to a DC voltage. An ADC is therefore a key component that can be used by the microcontroller to read values from all three of the chemical sensors.

The secondary target application is a nitric oxide (NO) releasing catheter. The NO released by the catheter prevents biofilm formation and increases the safety of a long-term implantable catheter. The microsystem must use a digital-to-analog converter to drive a copper electrode with a specific waveform to support studies of the optimal release of NO. An ADC allows the microcontroller to monitor and report on the operation of its own DAC, which is essential in a prototype system.

With the requirement of a single-chip solution, an off-the-shelf component is not possible, so a custom ADC must be tailored to the unique requirements of these applications.

## 5.2 Desired Specifications

The NO releasing catheter requires a digital-to-analog converter to drive a copper electrode to 1.5 V, so the digital-to-analog converter is designed to operate at the 1.8 V rail. The ADC must be able to verify and calibrate the digital-to-analog converter, so it must also operate at the 1.8 V rail. This requirement is the first of several significant specifications that guided the design of the ADC.

The target applications for this microsystem share two challenging requirements: small volume and low-power. To minimize volume, all circuitry components are integrated onto a single chip in a 65 nanometer process. Each circuit is custom-designed to minimize power consumption and maximize battery life while meeting other design constraints.

The accuracy requirement of the ADC can be derived from the accuracy requirements of each of the sensors, while building in margin to increase the likelihood of success. The accuracy requirements for each sensor are given in Table 5.1.

Sensor characteristics can be used to translate each of these sensor specifications into ADC specifications. A polymeric membrane pH sensor available from e-SENS varies by 59 mV/pH [86], so to achieve a resolution of 0.05 pH, the ADC must have a voltage resolution of:

$$59 \text{ mV/pH} \cdot 0.05 \text{ pH} = 2.95 \text{ mV} \quad (5.1)$$

This required voltage resolution can then be converted into the necessary ADC resolution by dividing the full-scale ADC voltage (1.8 V) by the required resolution and taking a logarithm:

$$\log_2 \left( \frac{1.8 \text{ V}}{2.95 \text{ mV}} \right) = 9.25 \text{ bits} \quad (5.2)$$

To build in margin, 12-bit resolution is targeted. While a relatively simple ADC architecture, such as a standard successive-approximation ADC, can achieve 8 or perhaps

**Table 5.1** Sensor data and time resolution requirements.

Requirement	Specification
pH Sensitivity	0.05 pH or less
pH Range	4 to 7
Conductivity Range	.05 S/m to 1.8 S/m
Conductivity Resolution	.1 S/m
Conductivity Frequency	1 kHz
Temperature Range	10 °C to 45 °C
Temperature Resolution	0.2 °C

10 bits of resolution, to achieve 11 or 12 bits requires more sophisticated design techniques, such as postfabrication calibrated trimming of capacitors or a more complex architecture that is more immune to process variation such as a sigma-delta architecture.

To form a conductivity sensor, an operational transconductance amplifier (OTA) is used to convert an AC voltage provided by a DAC into an AC current. This AC current is then applied to the conductivity sensor array, which acts as a resistance and presents an AC voltage to the ADC, as shown in Figure 5.1. The value of the resistor in Figure 5.1 should be chosen to limit the maximum current draw of the conductivity sensor and to keep the maximum current draw within the capabilities of the opamp. The opamp's maximum current output, listed in Table 4.2, is 1.65 mA. To create margin, the maximum targeted current drawn of the conductivity sensor is selected as 450  $\mu$ A. The value of the feedback resistor is

$$R = \frac{V_{DD} - V_{com}}{450 \mu A} = \frac{0.9 V}{450 \mu A} = 2 k\Omega \quad (5.3)$$

The resistance of the conductivity sensor ( $R_c$ ) is inversely proportional to the conductivity of the analyte ( $\rho$ ), with a conversion constant ( $k$ ) that can be set by the size and position of the electrodes. The conversion constant for the commercial conductivity sensor from e-SENS is  $\sim 400 \Omega S/m$ . When the maximum conductivity (1.8 S/m) is present,  $V_{DAC}$  presents a full 1.8 V peak-to-peak voltage, and the resulting waveform present at  $V_{ADC}$  is

$$V_{ADC} = \frac{k \cdot V_{DAC}}{R \cdot \rho} = 200 \text{ mV} \quad (5.4)$$

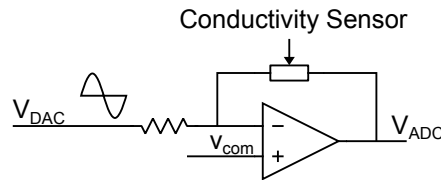
To determine the resolution needed to resolve 0.1 S/m, the voltage change that occurs in  $V_{ADC}$  when the conductivity is decreased to 1.7 S/m will be calculated. In this case, the voltage present at  $V_{ADC}$  is increased to

$$V_{ADC} = \frac{k \cdot V_{DAC}}{R \cdot \rho} = 212 \text{ mV} \quad (5.5)$$

Therefore the ADC must detect a change of 12 mV, which requires a resolution of

$$\log_2 \left( \frac{1.8 V}{12 \text{ mV}} \right) = 7.26 \text{ bits} \quad (5.6)$$

which is much less than the pH sensor requirement. The resolution requirement should also be checked when the conductivity is at a minimum. The microcontroller constantly checks



**Figure 5.1** Diagram of the conductivity sensor.

that  $V_{ADC}$  has not saturated, and lowers  $V_{DAC}$  if necessary to keep  $V_{ADC}$  near but below 1.8 V. When the conductivity is at 0.05 S/m,  $R_c$  is 8 k $\Omega$ , the OTA gain is 8 k $\Omega$ /2 k $\Omega$  = 4, and  $V_{DAC}$  is set by the microcontroller just below 1.8 V/4 = 450 mV peak-to-peak. When conductivity is 0.15 S/m, this value increases to 1.35 V, so low conductivity values do not increase the resolution requirement of the ADC. However, the conductivity sensor also requires a bandwidth of 1 kHz, which requires a finer time resolution than that required by the pH sensor.

For the temperature sensor, a DC current source will convert the sensor's electrical resistance to an analog voltage. The electrical resistance and therefore voltage increases with increasing temperature according to the following equation:

$$V = V_0[1 + \alpha(T - T_0)] \quad (5.7)$$

In this equation,  $V_0$  represents a starting voltage level measured at temperature  $T_0$  and  $V$  represents the resulting voltage at temperature  $T$ . The coefficient  $\alpha$  is the temperature coefficient of resistance, which for platinum is 0.00392/ $^{\circ}\text{C}$ . The temperature sensor on the sensor chip produced by e-SENS has a nominal resistance of 640  $\Omega$ . This resistance would lead to higher than desirable power dissipation, but in a final system the chip could be slightly customized to increase the nominal resistance and reduce power dissipation. If the resistance is tuned so that the maximum temperature (45  $^{\circ}\text{C}$ ) corresponds to the maximum voltage (1.8 V), then the minimum voltage will correspond to:

$$1.8 \text{ V} [1 + 0.00392/^{\circ}\text{C} \cdot (10^{\circ}\text{C} - 45^{\circ}\text{C})] = 1.55 \text{ V} \quad (5.8)$$

This is the voltage at which the temperature resolution corresponds to the smallest voltage resolution. Therefore, the voltage resolution required to achieve the desired temperature change of 0.2  $^{\circ}\text{C}$  is:

$$0.2^{\circ}\text{C} \cdot 0.00392/^{\circ}\text{C} \cdot 1.55 \text{ V} = 0.00122 \text{ V} \quad (5.9)$$

The required voltage resolution will once again be converted into a required ADC resolution:

$$\log_2 \left( \frac{1.8 \text{ V}}{0.00122 \text{ V}} \right) = 10.53 \text{ bits} \quad (5.10)$$

This is the most challenging resolution requirement of the three sensors, and the ADC should be designed to achieve the maximum requirements of all the three sensors, in terms of resolution and bandwidth.

The aforementioned derivations assume a full-scale voltage of 1.8 V (for example, see (5.2), (5.3), and (5.10)). Using the maximum available voltage will allow the ADC to

maximize dynamic range and achievable precision. Even though the core of the microsystem will operate at a lower voltage (1 V), 1.8 V will be available to satisfy the pad ring's requirements. The ADC must be able to verify the DAC's operation, and therefore must have an input range close to 1.5 V. To build in margin, it will use the 1.8 V rail and support a full rail-to-rail input. The ADC's requirements are summarized in Table 5.2.

## 5.3 ADC Design

### 5.3.1 Architecture Selection

The first step in designing an ADC after determining the desired specifications is to survey the available ADC architectures and determine which architecture is most appropriate for the specifications. Common analog-to-digital architectures include flash, folding, successive approximation register (SAR), pipelined and sigma-delta. The relative advantages and the specifications for which each ADC type is suited were evaluated based on widely accepted figures-of-merit. The first is

$$P = 2^B \cdot f_s \quad (5.11)$$

where  $B$  is the effective number of bits and  $f_s$  is the sampling frequency. Figure-of-merit  $P$  therefore considers only the accuracy and speed of the ADC. Figure-of-merit  $F$  adds consideration of power:

$$F = \frac{2^B \cdot f_s}{P_{diss}} \quad (5.12)$$

where  $P_{diss}$  represents power dissipation. In power-constrained applications, such as the applications examined by this dissertation, figure-of-merit  $F$  is a more appropriate guide to architectures. Surveys of recent ADCs are available in two review articles, which collectively report the figures-of-merit of over 900 commercially-available ADCs [87, 88].

**Table 5.2** ADC requirements.

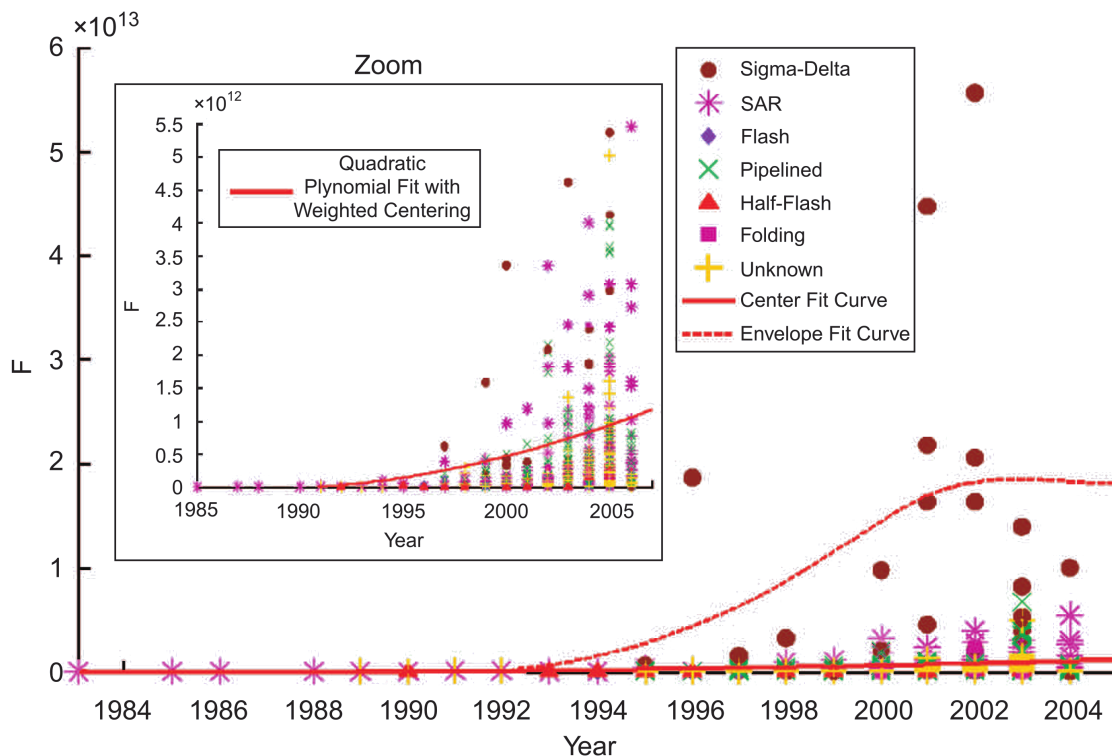
Specification	Requirement
Accuracy	11.6 bits
Bandwidth	1 kHz
Voltage range	0 to 1.8 V (rail-to-rail)
Supply voltage	1.8 V
Power	As low as possible
Sleep mode	Required
Component matching tolerance	As high as possible

While pipeline ADCs achieve the highest performance, as demonstrated by their high  $P$  figure-of-merit, the sigma-delta ADCs are best suited for high-precision and low-bandwidth applications. Sigma-delta ADCs consistently achieve the highest figure-of-merit  $F$ , as shown in Figure 5.2.

Other ADC designs fit other parts of the design space. Successive-approximation ADCs are the most traditional and popular type, and are suitable for many ADCs with balanced, moderate requirements. Folding ADCs often have the best performance-to-cost ratio. Flash ADCs are best suited for high-bandwidth and low-precision applications.

One review concludes that sigma-delta converters are advantageous when high resolution (at least 12 bits) is required, and low to medium bandwidth (kilohertz to megahertz) is acceptable [88]. This is even more the case when minimizing power is a high priority.

In addition to achieving high resolution and low power consumption, the sigma-delta architecture is also attractive because it is relatively forgiving with respect to component matching accuracy. Achieving high component matching in CMOS requires both expert layout and postfabrication calibration. An architecture that is tolerant of CMOS variability



**Figure 5.2** Figure of merit  $F$ , by ADC architecture. The sigma-delta architecture usually achieves the most power-efficient designs. Taken with permission from [87], ©2011 IEEE.

is preferred. The sigma-delta architecture achieves high resolution even without perfect device matching.

A common method of reducing power consumption is to employ circuits that operate at very low voltages [89,90]. Although very low-voltage circuits achieve the most minimal power levels, this design technique is not available for this project because the design requires a large input range of 1.8 V.

### 5.3.2 Low-power Sigma-delta Literature Review

Within the sigma-delta ADC family, there are a number of variants. Their key characteristics are summarized in Table 5.3. Some new designs eliminate the opamp in a sigma-delta stage, replacing it with a comparator or a zero-crossing detector [88]. This design is attractive because opamps are the dominant source of power consumption in a sigma-delta architecture, but without negative feedback their accuracy is diminished [91]. For example, a prototype comparator-based pipeline ADC achieved only 8.6 bits [92]. Another reported zero-crossing based ADC (though using a pipeline architecture) achieved only 8 bits [93]; although both of these designs achieve impressive frequency and power characteristics. The reported results indicate that while these designs are excellent for higher frequencies and lower resolutions, and can achieve lower power than their opamp-based equivalents, they have reduced precision because of overshooting in the charge transfer. The opamp's feedback prevents overshooting problems and achieves a more accurate charge transfer, limited only by its gain.

Another replacement for the opamp is an inverter [94]. The inverter includes negative feedback like the opamp, but the performance and power consumption are strongly dependent on the supply voltage, so a very precise regulator would be required to maintain a sufficiently accurate supply voltage. An opamp-based design has a better power supply

**Table 5.3** Characteristics of several example low-power ADCs. Key characteristics include the order and oversampling ratio (OSR) for sigma-delta ADCs, effective number of bits (ENOB), and power consumption.

Ref	Circuit technique	order	OSR	ENOB	Power	$f_s$
[92]	Comparator-based	Pipeline		8.6 bits	2.5 mW	7.9 MHz
[93]	Comparator-based	Pipeline		6.9 bits	4.5 mW	100 MHz
[94]	Inverter-based	3	42	10.5 bits	0.73 $\mu$ W	240 Hz
[95]	Double-sampling	3	128	15 bits	38 $\mu$ W	16 kHz
[96]	Single-opamp	3	50 to 125	12 bits	200 nW	90 Hz

rejection ratio and more consistent current consumption across a range of voltages, which is important in a battery-powered application.

There are several opamp-based designs that are closer to the target specifications given in Table 5.2. An opamp-based sigma-delta modulator designed for a hearing aid employs a 3<sup>rd</sup>-order architecture with an oversampling rate (OSR) of 128 [95]. Another example of a very low-power opamp-based sigma-delta ADC uses a single opamp to power a 3<sup>rd</sup>-order ADC by cleverly switching the opamp to integrate each of the three stages serially [96]. Although the OSR was not specified, given the sampling frequency and bandwidth specifications, it is between 50 and 125.

These designs illustrate that sigma-delta modulators that achieved the desired specifications of this project generally were based on opamps, were relatively low-order (either 2<sup>nd</sup>- or 3<sup>rd</sup>-order) to minimize the number of opamps, and had an OSR in the range of 64 to 128. Because the opamps in the design use the majority of the ADC's power consumption, an important key to reducing the ADC's power consumption is to design its opamps to minimize power consumption [97].

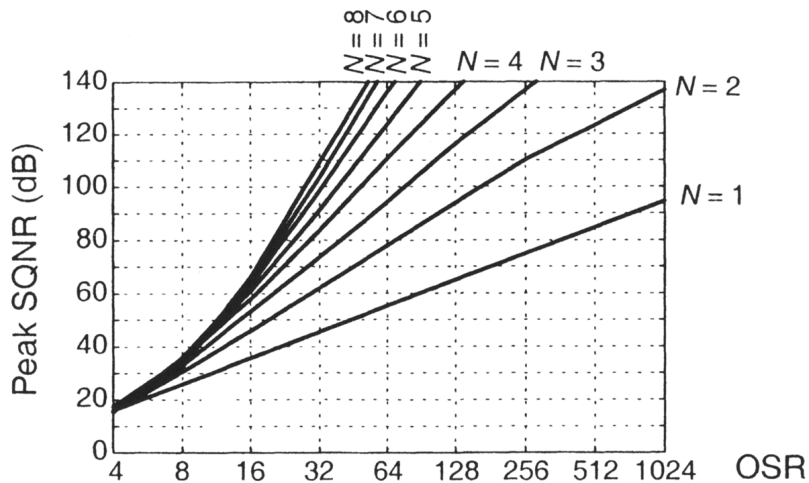
### 5.3.3 Sigma-delta Modulator Design

The highest level parameters must be determined first. For the design of a sigma-delta ADC, the highest level and most important parameters to be determined are the ADC order, the number of quantizer levels, and the OSR. Section 5.3.2 indicated that most low-power sigma-delta ADCs were of relatively low-order (2<sup>nd</sup>- or 3<sup>rd</sup>-order) and used only a 1-bit quantizer. These design decisions minimize the number of analog circuit devices, helping to reduce power.

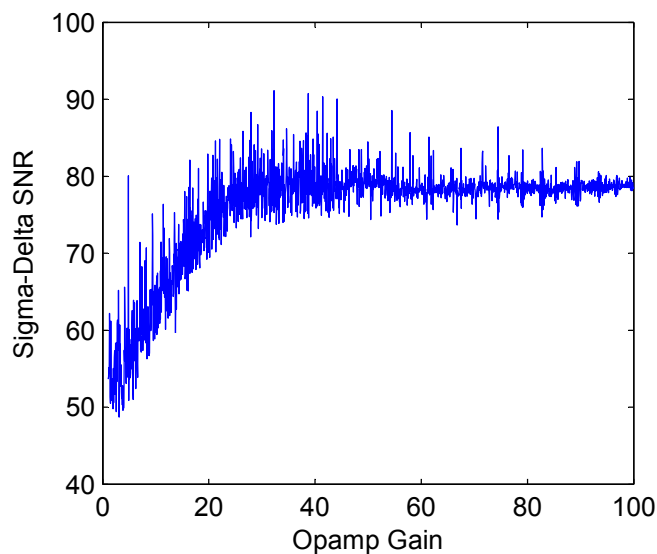
A well-known sigma-delta textbook includes a figure (reproduced here in Figure 5.3) which helps map the sigma-delta order and OSR to a maximum achievable precision [98]. That textbook indicates that a 2<sup>nd</sup>-order sigma-delta with a 1-bit quantizer and OSR of 128 achieves 94.2 dB, over 15 ENOB (equivalent number of bits). The authors of [98] say that they used “extensive computations” to generate Figure 5.3, which shows the “achievable peak signal-to-quantization-noise ratio (SQNR).” Varying different parameters of the simulation can result in considerable variation in the apparent SNR, as shown in Figure 5.4. However, our own simulations using ideal difference equations typically achieve only about 77.9 dB SNR. While the typical SNR value is below the achievable peak, it is sufficient to meet the application requirements.

The block diagram given in Figure 5.5 depicts a common topology for a 2<sup>nd</sup>-order sigma-delta, based on the Boser-Wooley design [99]. The optimal coefficients for a conventional

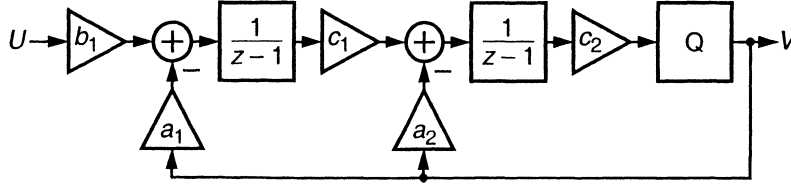




**Figure 5.3** Peak signal to quantization noise ratio (SQNR) by oversampling ratio (OSR). Peak achievable SQNR is based on simulations. Taken with permission from [98] ©2004, John Wiley and Sons.



**Figure 5.4** Sigma-delta SNR versus gain of the integrating opamp. This simulation used ideal difference equations and does not include circuit nonidealities. It demonstrates 1) that there is considerable variation in measured SNR from simulation to simulation, as discussed in Section 5.3.3 and 2) that the opamp needs to achieve only about 40 dB, as discussed in Section 5.3.3.2.



**Figure 5.5** Block diagram of a Boser-Wooley 2<sup>nd</sup>-order sigma-delta modulator. Taken with permission from [98] ©2004, John Wiley and Sons.

2<sup>nd</sup>-order sigma-delta are  $a_1 = b_1 = a_2 = \frac{1}{4}$  and  $c_1 = \frac{1}{3}$  [98]. The  $c_2$  coefficient is not significant in a 1-bit quantizer system because its output is fed directly into a comparator, so the output of the comparator will be the same regardless of the value of  $c_2$ .

Although a differential input is more typical for sigma-delta data converters, a single-ended design is used in this project for two reasons. First, a single-ended sigma-delta eliminates the need for any complex common-mode feedback circuitry, which would have increased design time and effort and increased power consumption. Second, a single-ended opamp is already needed as a unity gain input buffer for the pH sensor and as a buffer capable of driving many pF off-chip from the DAC. To reduce design time, this single-ended opamp is reused in a single-ended sigma-delta ADC design.

### 5.3.3.1 Capacitor sizing

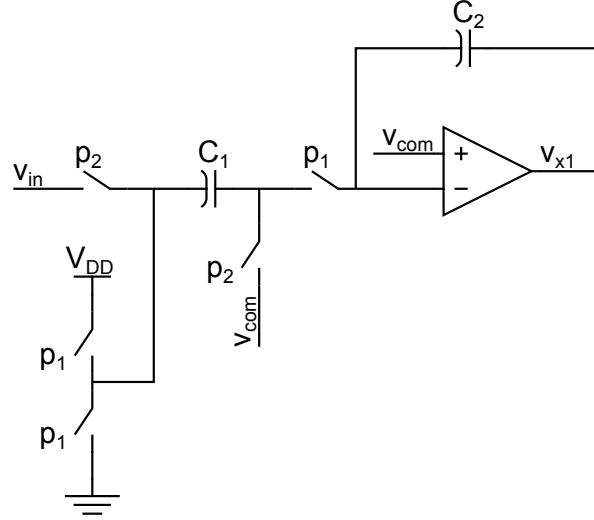
The block diagram must be translated into circuit elements before devices can be sized. For a single-ended design, the circuit elements necessary to implement the first stage of the block diagram in Figure 5.5 are depicted in Figure 5.6.

The parameters of the block diagram can be translated into real component sizes by relating the ideal difference equations that describe Figure 5.5 to charge transfer functions that describe Figure 5.6. The difference equation implemented by the first stage of the block diagram in Figure 5.5 is:

$$x_1(n+1) = x_1(n) + b_1 \cdot u(n) - a_1 \cdot v(n) \quad (5.13)$$

To find the capacitor ratios necessary to implement the desired difference equations, the variables in those difference equations can be related to voltage levels in the circuit. In the Matlab program which generated the optimal transfer coefficients, the variables  $x_1$ ,  $u$ , and  $v$  are all defined as having a range  $[-1, 1]$ . The equivalent voltages have the range  $[0 \text{ V}, 1.8 \text{ V}]$ . As an example,  $v_{x_1}$  is related to  $x_1$ :

$$x_1 = \frac{v_{x_1} - 0.9 \text{ V}}{0.9 \text{ V}} \quad (5.14)$$



**Figure 5.6** Circuit schematic of the first stage of the sigma-delta modulator. Node voltages are annotated, as well as driving clocks on switches.

While  $v$  is either -1 or 1, the digital feedback signal  $v_d$  is either 0 or 1. The voltage transfer equations can be derived from the charge induced on the capacitors during each clock cycle. During phase 2, the input  $u$  charges the capacitor relative to 0.9 V (held on the right side of the capacitor):

$$Q_u = (v_u - v_{com}) \cdot C_1 \quad (5.15)$$

For this entire charge to be transferred to the second capacitor, the input would then be held at  $v_{com}$  (0.9 V). Instead, it is moved to either VDD or ground (0 V) based on feedback from  $v_d$ , inducing an additional charge:

$$Q_{v_d} = -V_{DD} \cdot (v_d - v_{com}) \cdot C_1 \quad (5.16)$$

These two charges are next transferred to the second capacitor via the opamp, which effectively adds them to the charge already stored on  $C_2$ , whose voltage is referred to as  $v_{x1}$ :

$$Q_{v_{x1}}(n+1) = Q_{v_{x1}}(n) + Q_u + Q_{v_d} = Q_{v_{x1}}(n) + (v_u - V_{DD} \cdot v_d) \cdot C_1 \quad (5.17)$$

To convert the charges to voltages, the entire equation can be divided by the capacitance:

$$v_{x1}(n+1) = v_{x1}(n) + (v_u - V_{DD} \cdot v_d) \cdot \frac{C_1}{C_2} \quad (5.18)$$

After replacing voltages with their unitless equivalents (for example  $v_{x1}$  with  $x_1 \cdot 0.9 \text{ V} + 0.9 \text{ V}$  and  $v_d$  with  $(v+1)/2$ ) and then dividing by the voltage and simplifying, this is the resulting equation:

$$x_1(n+1) = x_1(n) + \frac{C_1}{C_2} \cdot u(n) - \frac{C_1}{C_2} \cdot v(n) \quad (5.19)$$

To make this equivalent to the desired difference equation (5.13), the capacitor ratio  $\frac{C_1}{C_2}$  must be set to the desired transfer coefficient: in this case,  $\frac{1}{4}$ . A similar procedure is used to find the capacitor ratios for the second stage. A complete (though simplified) schematic is given in Figure 5.7.

Noise considerations determine the actual sizes of the capacitors. The input capacitor must be large enough to ensure that its thermal noise (or  $kT/C$  noise) is less than half the expected quantization noise, so that its noise will not be the limiting factor in system performance [98]. Therefore, its noise contribution should be 3 dB lower than quantization noise, or less than -97 dB, relative to an input signal that is .9 V peak-to-peak. The power of the input signal is  $(.45 \text{ V})^2$  or approximately  $0.2 \text{ V}^2$ . Therefore, the allowable thermal noise contribution of the capacitor is:

$$v_n = \sqrt{\frac{0.2}{10^{97/10}}} = 6.3 \mu\text{V} \quad (5.20)$$

The equation for the thermal noise (or  $kTC$  noise) of a capacitor is:

$$\bar{v}_n^2 = \frac{2kT}{C} \quad (5.21)$$

The normal equation is doubled because noise can be accumulated from both the charging and discharging events that occur in one clock period. Given that capacitor noise, the minimum size of the input capacitor is:

$$C_1 = \frac{2kT}{v_n^2 \cdot \text{OSR}} = 1.6 \text{ pF} \quad (5.22)$$

Later simulations showed that charge-injection from the switches was reducing SNR. The unusually large charge-injection was due in part to the use of I/O transistors to facilitate a full 1.8 V input. These I/O transistors have a relatively large minimum area and therefore relatively large minimum capacitance compared to the 1.2 V transistors. To mitigate the effects of the charge injection, all switches were implemented as pass-gates, with an NMOS and PMOS transistor in parallel, and the sizes of the capacitors were increased. For example, the input capacitor was increased to 3 pF.

As derived above, the size of the second capacitor,  $C_2$ , is four times the size of  $C_1$ , or 12 pF. The opamp used in each integration stage is identical, to maximize circuit reuse and minimize design time, so it must be equally capable of driving a large capacitance as the opamp in the first stage is capable of driving its capacitance. However, to reduce area, the integrating capacitance of the second stage is reduced by half to 6 pF. This capacitor size reduction does not increase noise because the noise introduced by the second stage is shifted into the high frequencies by the ADC's feedback mechanism.



### 5.3.3.2 Opamp requirements

Some opamp specifications can be derived from the ADC's requirements. To ensure that the charge transferred each clock cycle is accurate, only 25% of each half-clock-period should be devoted to slewing, so that the rest of the clock period can be used to allow the opamp voltage to accurately settle to its final value. Equivalently, enough current must be provided to charge the integrating capacitor in  $1/8$  of the clock period, or 8 times the amount of current that would be necessary to charge the capacitor in one clock period:

$$I = 8 \cdot f_{CK} \cdot C_1 \cdot V_{DD} = 8.8 \mu\text{A} \quad (5.23)$$

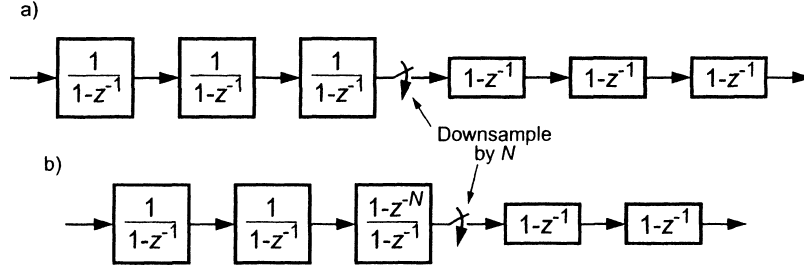
There is no straightforward way to calculate the required gain of the opamp, so a Matlab simulation implementing the desired difference equations was modified to account for finite opamp gain to test the effect of opamp gain on sigma-delta performance. Although this method of simulation assumed perfect linearity, the gain requirements of the opamp were surprisingly modest.

As shown in Figure 5.4, after an opamp gain of only 40 dB, gain no longer has a significant effect on the sigma-delta SNR. Opamp linearity has a much more dramatic effect on sigma-delta SNR, but simulating opamp nonlinearity is much more difficult, as is improving an opamp's linearity.

### 5.3.4 Decimation Filter Design

A common decimation filter is a 3<sup>rd</sup>-order sinc filter (depicted in Figure 5.8), which is highly power and computationally efficient but lacks the precision necessary for full decimation. It is often followed by a high-precision half-band FIR (finite impulse response) filter to complete decimation. Bit-accurate simulations of both of these filters were completed in Matlab, which allows very fast simulations. The 3<sup>rd</sup>-order sinc filter (also called a cascaded interleaved comb filter) was initially implemented in Matlab, and checked against a Matlab implementation written by Shen-Hsiao Tan [101]. The final Verilog implementation also exactly matched both Matlab filters.

The Delta Sigma Toolbox aided the rapid development of the decimation filter [102]. The toolbox includes a function to generate a half-band FIR filter based on two parameters: the required passband frequency and the absolute value of the passband and stopband ripple. The function generates the half-band filter based on the method originally developed by Saramaki [103]. The parameters were relaxed as much as possible to minimize hardware complexity while still achieving 11-bit resolution. The eventual filter used only  $3 \cdot 3 + 5 \cdot (3 \cdot 4 + 4 - 1) = 84$  additions per output sample, at a very low clock speed.



**Figure 5.8** Hogenhauer form of a  $\text{sinc}^3$  filter. Subfigures are a) the  $\text{sinc}^3$  decimation filter, and b) the more efficient Haugenhauer form [100] of the same filter. Taken with permission from [98] ©2004, John Wiley and Sons.

A mathematical model of the half-band FIR filter was used to verify performance in Matlab. A second Matlab implementation modeled actual hardware structures and matched the initial Matlab version. Finally, a Verilog implementation matched both Matlab versions.

## 5.4 Results

### 5.4.1 Simulation Parameter Selection

Although circuit-level simulations are the best approximations of postfabrication silicon performance, these simulations can only be completed after a substantial amount of circuit design. Initial design was performed in Matlab, because Matlab simulations can be executed quickly, and designs can be verified and optimized quickly. After prototyping in Matlab, circuit-level simulations in Cadence using the Spectre circuit simulator resulted in more accurate simulations and were used to refine the modulator design. Synopsys NC-Verilog simulations were only used to verify that Verilog filter implementations match their Matlab counterparts, since these digital circuits match their models exactly.

Testing parameters must meet several important constraints to maximize the achieved signal-to-noise ratio:

- The inverse of the sampling frequency must be a round number (such as 100 ps), so that the pulse length can be exactly specified in Cadence.
- To be a coherent simulation, the input frequency must fall exactly into an FFT bin. That is, the input frequency multiplied by the simulation time must be an integer.
- The number of output samples (sampling frequency  $\times$  simulation time) must be exactly divisible by the OSR, which is 128, if the decimation filter is also simulated.

These constraints ensure that the simulation is coherent. A large input amplitude also helps to maximize SNR. The Hann window function spreads the input sine wave's power

into 3 FFT (Fast Fourier Transform) bins, but ensures that noise is not aliased into the passband when the FFT is taken to determine SNR. The simulation parameters listed in Table 5.4 meet these constraints.

#### 5.4.2 Simulated Performance of the Modulator

The Cadence simulation of the modulator circuit achieves, neglecting a small input offset, 88.6 dB SNR, which is an equivalent of 14.4 bits of resolution, using an input amplitude that is  $2/3$  of maximum input amplitude. The output spectrum of the simulation is shown in Figure 5.9. The circuit consumes only 222  $\mu\text{W}$  at 1.8 V.

This performance appears to exceed difference equation simulations used to explore expected performance. Matlab simulation of the difference equations, with an input sine wave  $2/3$  of maximum, achieved only 82.4 dB SNR. However, as shown in Figure 5.4, the achieved SNR can vary significantly from simulation to simulation. The 88.6 dB SNR neglects the DC bin, because the noise in that bin is due to a slight DC offset that is difficult to correct for in circuit simulations.

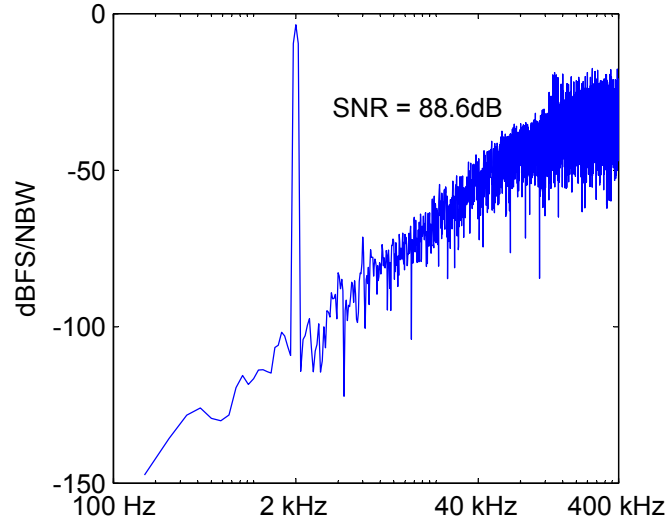
#### 5.4.3 Simulated Performance of the Decimation Filter

Matlab simulations of the decimation filter indicate that relatively little accuracy is lost by the filter, even while using a fairly minimal number of adders. The 3<sup>rd</sup>-order sinc filter induces some droop in the passband, but the ADC is designed to provide more bandwidth than is (omitting “strictly”) necessary for the application requirements. The input sine wave is 2 kHz, and is mostly unaffected by this droop. There is a slight decrease in SNR after the decimation filter due to noise that is folded into the passband by decimation. In the case of the output of the Cadence circuit simulation of the modulator, the resulting performance after the output is fed into the decimation filters is 79.6 dB SNR, which is lower than the 88.6 dB SNR present before the decimation filter. The final frequency spectrum is shown in Figure 5.10.

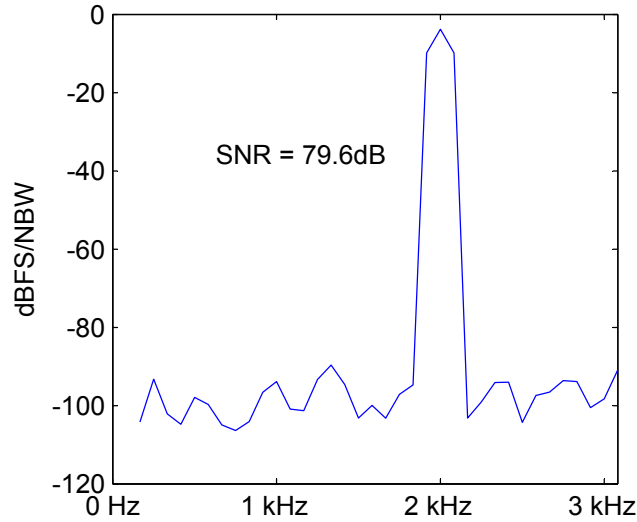
**Table 5.4** ADC simulation parameters.

Sampling frequency	800 kHz
Input frequency	2 kHz
Simulation time	12.8 ms
Input amplitude	600 mV ( $2/3$ of maximum)
Windowing function	Hann





**Figure 5.9** Circuit simulation of modulator showing 88.6 dB SNR.



**Figure 5.10** Frequency spectrum after the decimation filter.

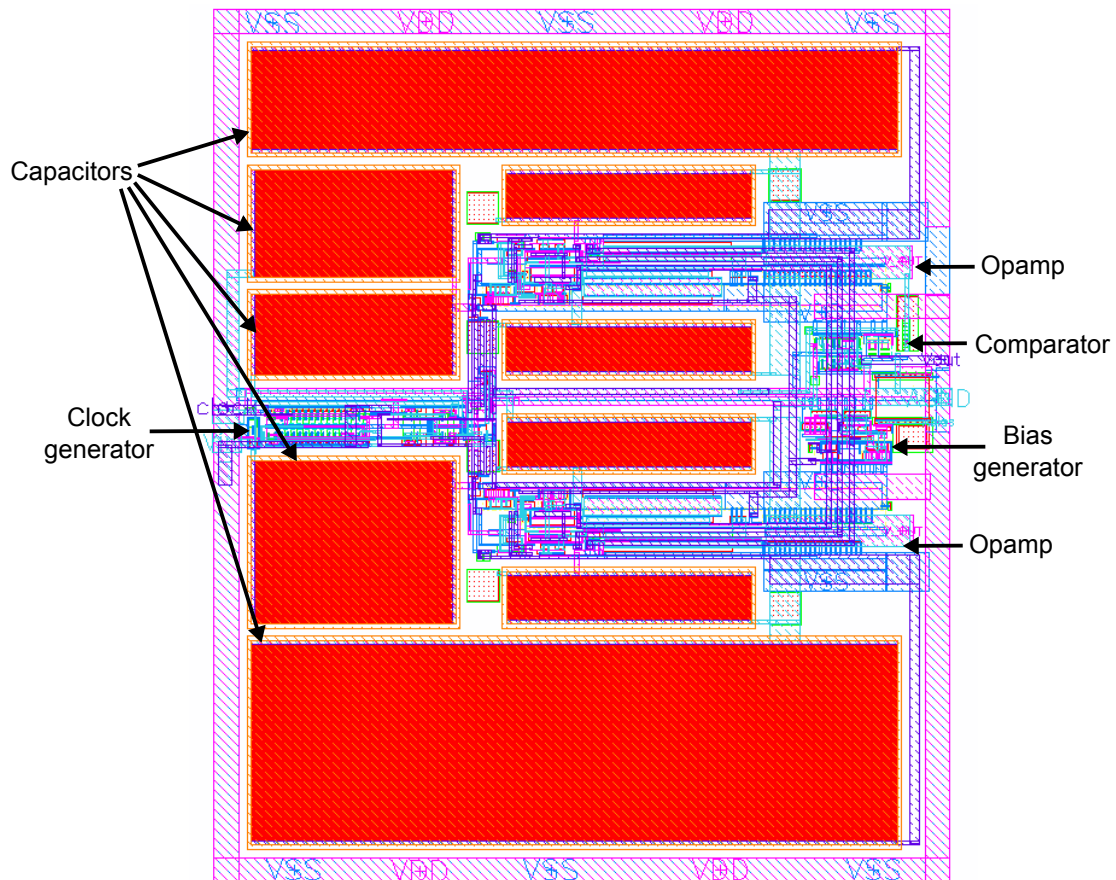
Synopsys Design Compiler estimates that the decimation filter, at a 1 MHz sampling rate, will consume 1.04  $\mu\text{W}$  dynamic power, plus 89 nW leakage power. Lower sampling frequencies will result in approximately proportionally lower dynamic power. Simulation results are summarized in Table 5.5.

#### 5.4.4 Measured Results

The fabricated ADC is depicted in Figure 5.11. The Cadence circuit simulations do not take into account device noise, so the SNR in those simulations are primarily limited by quantization noise. If device or environmental noise exceeds quantization noise, then the

**Table 5.5** ADC simulation results.

Modulator SNR	82.9 dB SNR
Modulator power consumption	222 $\mu$ W at 1.8 V
Decimation filter power consumption	1 $\mu$ W
SNR after decimation filter	81.1 dB SNR

**Figure 5.11** Layout of the fabricated sigma-delta modulator.

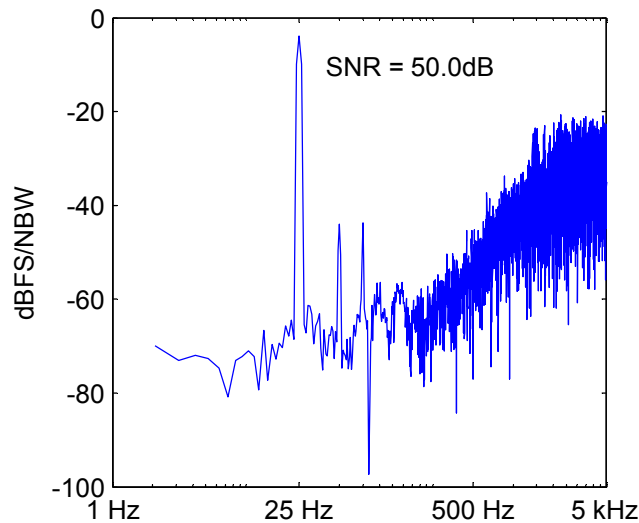
measured SNR is below the simulated SNR.

The performance of the ADC was greatly affected by the same issues that affected the opamp, which are described in Section 4.4. To summarize, a design error combined with layout effects resulted in a reduction of the opamp's current draw. This in turn reduced the opamp's slew rate and bandwidth. Based on the noise analysis in Section 4.3.5, reduced bias current also increases the thermal noise of the opamp. Because the opamp is used in the ADC's integrators, its increased noise reduces the ADC's SNR. The benefit of this error is that the ADC power consumption is reduced to  $2.7 \mu\text{W}$ .

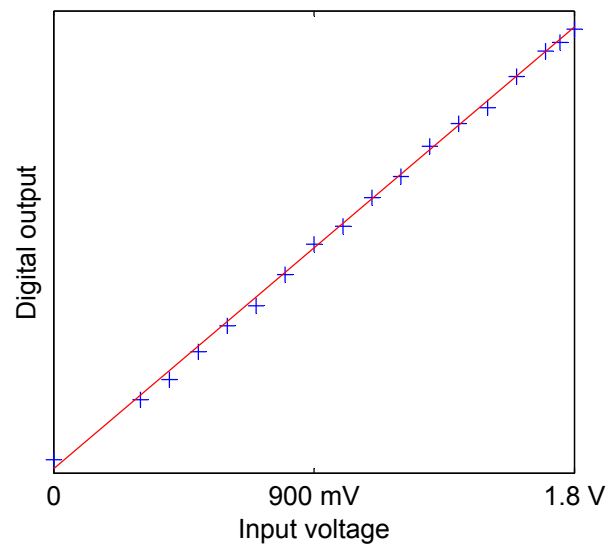
The output of the ADC was contaminated by the input signal, indicating that a lack of shielding led to signal isolation problems, which may have also contributed to a reduced measured SNR. The peak measured SNR is 50 dB, or 8 bits, as depicted in Figure 5.12. The harmonics apparent at 50 Hz and 75 Hz may be the result of input signal contamination of the integrator voltages.

The ADC was also subjected to DC inputs to examine its linearity. This was done to verify the operation of the ADC, since SNR figures were not very repeatable, and because the target applications often require DC values to be read. For each DC input, 8 16-bit output samples were averaged together. This could easily be done in software by the microcontroller. The result is in Figure 5.13, which compares input voltage to output code.

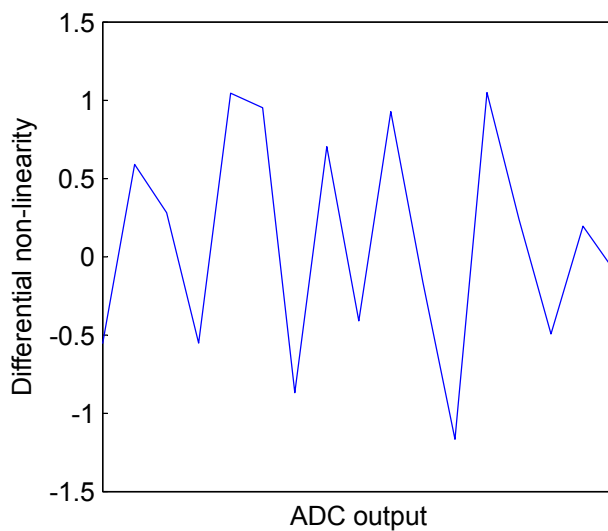
The differential and integral nonlinearity are depicted in Figures 5.14 and 5.15, respectively. The LSBs on the y axis of each of these figures are calculated assuming a 6-bit ADC. The max DNL is approximately 1 LSB, with the max INL slightly larger than 1 LSB. This level of performance corresponds approximately with a 6-bit ADC. Although the ADC



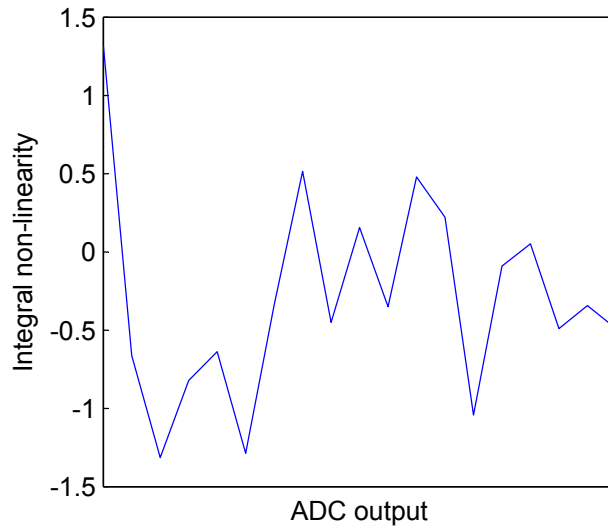
**Figure 5.12** Sigma-delta modulator measurement showing 50.0 dB SNR.



**Figure 5.13** Sigma-delta modulator measured DC linearity plot. For each data point, a DC voltage was input, the modulator was clocked at 1 kHz, and 8 outputs of the decimation filter were averaged to form one output in the figure. Two data points were discarded due to faulty data.



**Figure 5.14** Differential nonlinearity of ADC output. The y-axis is in LSBs of a 6-bit ADC.



**Figure 5.15** Integral nonlinearity of ADC output. The y-axis is in LSBs of a 6-bit ADC.

was able to achieve 8-bit resolution based on Figure 5.12, that result was the maximum of several measurements, whereas the linearity results were performed only once, and represent typical performance.

#### 5.4.5 Result Summary

The simulation results are summarized in Table 5.5, and the measured results in Table 5.6. In simulation, the modulator achieves the desired precision with very low operational power. It also illustrates substantial design reuse of the opamp, which is used in many system components.

The measured results indicate that the fabricated ADC does not achieve the desired specifications. The fabricated ADC does, however, achieve a peak SNR of 50 dB, or 8-bit resolution, and the linearity tests indicate that typical resolution is approximately 6 bits.

### 5.5 Conclusion

All common ADC architectures are considered for the unique requirements of the target applications, and the sigma-delta architecture is found to be the most suitable. Each step

**Table 5.6** ADC measured results.

Modulator SNR	50 dB SNR
Modulator power consumption	2.7 $\mu$ W at 1.8 V
DC input accuracy	6 bits

in the design process is explained, illustrating clearly how the design is optimized for its specifications. This design reuses circuits to an unusually high degree and is fabricated in a single-chip microsystem. In simulation, the ADC meets all specifications in spite of charge-injection from relatively large transistors, which are required to operate at 1.8 V in the 65 nm process. It also achieved an impressively low power level of 222  $\mu\text{W}$ .

While the fabricated ADC does not achieve its desired bandwidth or precision, the power level is also reduced to only 2.7  $\mu\text{W}$ . The decreased performance is primarily attributable to the decreased performance of the opamp. The reasons for this decreased performance have been identified in Section 4.4, and could be corrected in a future revision. In spite of this shortfall, this research makes several important contributions. The analysis in Sections 5.2 and 5.3 demonstrates how to select and customize an ADC architecture for low-power sensors. The simulation results show that the required specifications are achievable if opamp noise can be sufficiently reduced, and the integrated and fabricated chip shows that it is possible for researchers to design, integrate, and fabricate small low-power microsystems for the types of emerging biomedical applications targeted by this dissertation.

## CHAPTER 6

### A LOW-POWER SINGLE-OPAMP CHARGE SCALING 10-BIT DAC

#### 6.1 Introduction

The two primary target applications, the S-IVR and NO-releasing catheter discussed in Sections 1.4 and 1.5, both require a digital-to-analog converter (DAC). The S-IVR includes a conductivity sensor, which requires a DAC. The DAC's output is converted to a current which is driven across the conductivity sensor's electrodes. An ADC can then convert the resulting voltage into a digital conductivity measurement, as discussed in Section 5.2. The DAC is therefore essential for the S-IVR microsystem.

The NO-releasing catheter requires a DAC to drive a copper electrode with a specific waveform to release the NO. In this application, the DAC must be able to achieve high drive currents while minimizing its own power consumption.

With the requirement of a single-chip solution, a custom DAC must be tailored to the unique requirements of these applications.

#### 6.2 Desired Specifications

The target applications for this microsystem share two challenging requirements: small size and low-power. To minimize size, all circuitry components will be integrated onto a single chip in an advanced nanometer process. Each circuit will be custom-designed to minimize power consumption and maximize battery life while meeting other design constraints.

The NO releasing catheter requires a DAC to drive a copper electrode, which causes an electrochemical reaction that results in the release of nitric oxide (NO). Höfler et al. reported that the voltage waveform necessary for this reaction varies from  $-0.97$  V to  $-0.07$  V vs. a Ag/AgCl reference electrode [22]. From a circuit perspective, this could be accomplished by connecting the Ag/AgCl electrode to the 1.2 V rail, and then setting the DAC output to 0.23 V and 1.13 V, respectively. However, this configuration allows for very little margin

of error. If the DAC failed to achieve its full range by even a tenth of a volt, then it would be unable to drive the desired voltages. For a research system, a larger voltage range was desirable so that other electrical waveforms could be tested. Dr. Meyerhoff’s research group (who produced [22]) also communicated that a larger voltage range, up to 1.5 V, is desirable for testing, so the DAC must operate at the 1.8 V rail to accommodate this larger range.

The DAC must also be able to drive multiple mA of current, both for the conductivity sensor and potentially for the NO electrode. The conductivity sensor requires a resolution of 9.26 bits, as derived in (5.3). A DAC can only be designed to achieve an integer number of bits (except for sigma-delta DACs), so the resolution requirement can be rounded up to 10 bits. A constant draw of several mA would quickly deplete the system’s small battery, so it must only supply the high current when necessary, and operate at much lower power whenever possible. The accuracy requirements for the DAC are given in Table 6.1.

## 6.3 DAC Background

### 6.3.1 Common DAC Architectures

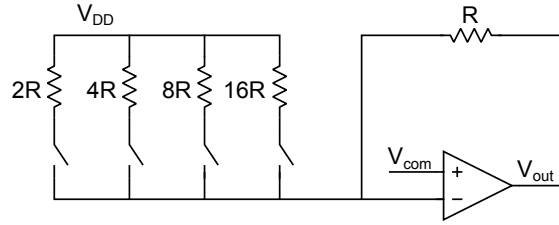
Most DACs can be categorized as current scaling, voltage scaling, or charge scaling. Each of these architectures will be discussed briefly.

An example current scaling DAC is shown in Figure 6.1. The DAC must operate in an environment with a single power rail and ground, although a reference voltage ( $V_{\text{com}}$ , usually  $V_{\text{DD}}/2$ ) may be generated. When one of the switches is closed, it creates a voltage drop of  $V_{\text{DD}}$  across one of the resistors, because the input to the negative terminal of the opamp is a virtual short to  $V_{\text{com}}$ . Whatever current is flowing into the negative terminal of the opamp from the parallel resistors, that same amount of current flows over the feedback resistor (labeled “R” in Figure 6.1), resulting in an output voltage at  $V_{\text{out}}$  between  $V_{\text{com}}$  and ground. This configuration is similar to an inverting summing amplifier. Allen and Holberg list the advantages of the current scaling DAC as being fast and insensitive to parasitic capacitance [104].

**Table 6.1** DAC specifications.

Requirement	Specification
Drive current	>10 mA
Operating power	As low as possible
Output voltage range	0 to 1.8 V
Resolution	10-bits





**Figure 6.1** Simplified schematic of a 4-bit current scaling DAC.

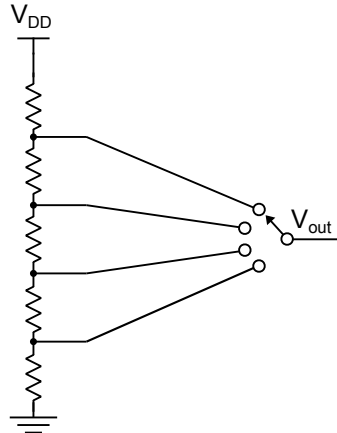
Although one disadvantage is that current is constantly flowing through the resistors, this current can be reduced by using larger resistor values. To keep the extra current less than that used by the opamp, which is  $\sim 60 \mu\text{A}$ , the resistor must be larger than

$$R = \frac{1.8 \text{ V}}{60 \mu\text{A}} = 30 \text{ k}\Omega \quad (6.1)$$

This size of resistor would result in a power consumption of approximately double the opamp alone. Another issue is that the range of  $V_{\text{out}}$  is constrained.  $V_{\text{out}}$  must be below  $V_{\text{com}}$ , which limits the output range of the current scaling DAC. The current scaling DAC cannot easily achieve a full-scale output range while using only a single opamp.

Another class of DACs is the voltage scaling DAC. A simplified schematic of this type of DAC is depicted in Figure 6.2. This DAC consists of resistors in series, with different voltage levels present in between each resistor. Switches are used to select which voltage level is desired.

This design overcomes the output range problem of the current scaling DACs: it natively achieves a full-range output. Only a single opamp would be required to provide a unity-gain buffer to the output. One drawback for this design is the sheer number of resistors needed for its implementation, which can be calculated as:



**Figure 6.2** Simplified schematic of a 2-bit voltage scaling DAC.

$$\text{resistors} = 2^{\text{bits}} + 1 \quad (6.2)$$

A 10-bit DAC requires over 1,000 resistors. Creating that many resistors would be impractical to lay out by hand, without automatic layout tools or scripts. The large number of resistors requires many switches to select the desired output voltage. Even with a mux tree, the switches can consume significant leakage current. The parasitic capacitance of the switches may slow the operation of the DAC, or distort the output voltage.

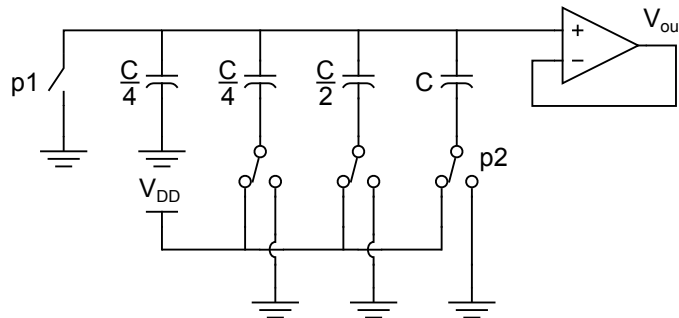
The last basic DAC architecture that will be considered here is a charge scaling architecture. A simplified schematic of a 3-bit charge scaling DAC is depicted in Figure 6.3. A charge scaling DAC uses capacitors rather than resistors, and therefore does not have any constant current flowing outside of leakage and the unity-gain buffer of the opamp.

In Figure 6.3, there are two nonoverlapping clocks, labeled p1 (which applies to only the left-most switch), p2, which applies to each of the switches connected to capacitors. During the first phase, referred to as p1, the switch controlled by p1 is closed and all of the p2 switches connect the capacitors to ground, so that both sides of the capacitors are held at ground and all of the capacitors are discharged. During the second phase, referred to as p2, the switch controlled by p1 is open and some of the switches triggered by p2 will switch from ground to  $V_{DD}$ , depending on the digital input word. The circuit effectively is transforming into two separate capacitors, connected in series, between  $V_{DD}$  and ground, as depicted in Figure 6.4.

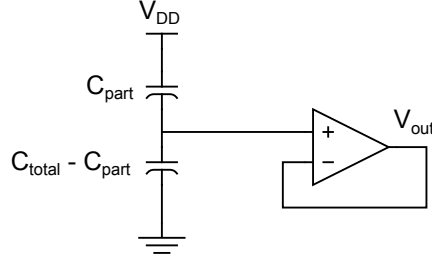
In Figure 6.4,  $C_{total}$  represents total capacitance, which, using the example in Figure 6.3, is  $2C$ .  $C_{part}$  is the portion of the capacitance that is connected to  $V_{DD}$  because of the digital input word. The equivalent capacitance of the two capacitors in series is:

$$C_{eq} = \frac{C_1 \cdot C_2}{C_1 + C_2} \quad (6.3)$$

The equation for finding the voltage across one of the capacitors ( $C_1$  in this example) in series is:



**Figure 6.3** Simplified schematic of a 3-bit charge scaling DAC.



**Figure 6.4** Simplified schematic of a charge scaling DAC during clock phase p2.  $C_{total}$  is the total capacitance. The capacitors connected to  $V_{DD}$  have been combined into  $C_{part}$ , and the remaining capacitors are connected to ground and have a capacitance of  $C_{total} - C_{part}$ .

$$V_1 = V_{eq} \cdot \frac{C_{eq}}{C_1} \quad (6.4)$$

$V_{eq}$  is the voltage across both capacitors. Applying these equations to the circuit shown in Figure 6.4, using  $C_{part}$  for  $C_2$  and  $C_{total} - C_{part}$  for  $C_1$ , we can plug these values into (6.3) to solve for the equivalent capacitance of the capacitors in series.

$$C_{eq} = \frac{C_{part} \cdot (C_{total} - C_{part})}{C_{total}} \quad (6.5)$$

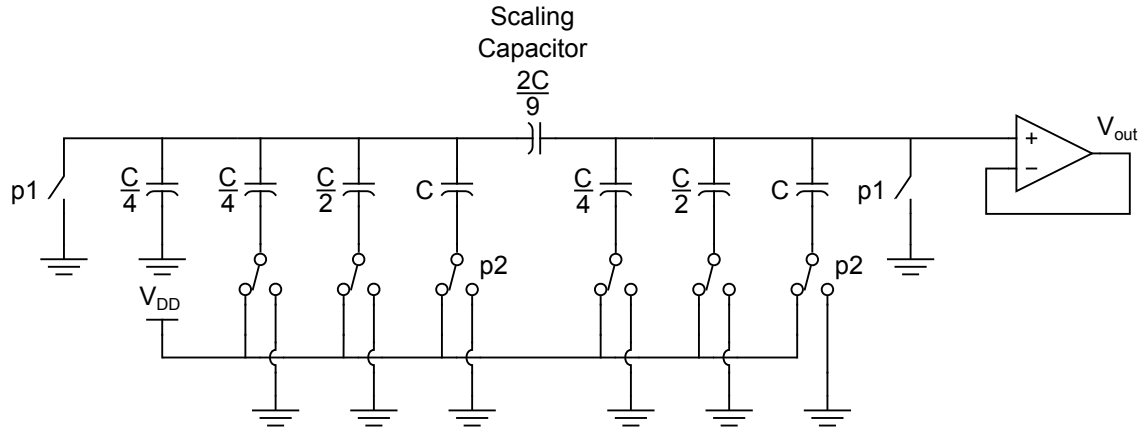
This result can then be inserted into (6.4) to find the voltage across the bottom capacitor,  $C_{total} - C_{part}$ :

$$V_{out} = V_{DD} \frac{C_{eq}}{C_{total} - C_{part}} = V_{DD} \frac{C_{part} \cdot (C_{total} - C_{part})}{C_{total} \cdot (C_{total} - C_{part})} = V_{DD} \frac{C_{part}}{C_{total}} \quad (6.6)$$

The result in (6.6) can be intuitively understood. The output voltage will be the same fraction of the capacitance that is connected to  $V_{DD}$  during the second phase of the clock.

The charge scaling DAC will be used for this design because it does not require very large area, has no DC current other than leakage and opamp current, and requires only one opamp that can act as a unity-gain buffer so that high loads do not distort the output voltage. The simple design depicted in Figure 6.3 can be used for low-precision DACs, but for a 10-bit DAC, the ratio between the capacitor used for the LSB (least significant bit) and the capacitor used for the MSB is too large. The 10-bit resolution cannot be achieved if components are not matched within  $0.5 \cdot \text{LSB}$ . Techniques have been developed to extend the resolution of the charge scaling DAC without requiring such a high disparity between sizes of capacitors used in the design.

An extended resolution charge scaling DAC is depicted in Figure 6.5. In this figure, the capacitors and switches to the left of the scaling capacitor correspond to the least significant bits, and the switches to the right of the scaling capacitor represent the most significant bits. As shown by (6.3), placing a capacitor in series with another capacitor reduces the



**Figure 6.5** Simplified schematic of an extended resolution 6-bit charge scaling DAC.

total capacitance of the capacitors in series. The scaling capacitor in Figure 6.5, from the perspective of the output voltage, does exactly this with the array of capacitors to its left. The scaling capacitor is sized so that it and all the capacitors to its left combine to act as a single capacitor the same size as the previous terminating capacitor it is replacing, which is  $C/4$  in the example of Figures 6.3 and 6.5.

### 6.3.2 Low-power DAC Literature Review

In general, the literature on low-power DACs is far more sparse than the literature on ADCs (analog-to-digital converters). This is due, in part, to the more simple nature of the basic DAC designs, as shown in Section 6.3.1. A recent textbook on DACs and ADCs included chapters on high-speed and high-resolution DACs, but no section on low-power DACs [105]. A few low-power DACs reported in the literature will be considered here.

A low-power low-throughput 6-bit DAC was presented in a textbook on low-power design [106]. This DAC achieved an impressively low  $440\ \mu\text{W}$  power consumption, and operated from a  $1.5\ \text{V}$  power supply, but output a reduced range of only  $0.7\ \text{V}$ . This DAC achieved such low power consumption in part by matching current drivers to a known external load. This convenience does not exist for the NO catheter application, in which the DAC may need to drive a variety of loads, potentially with high currents.

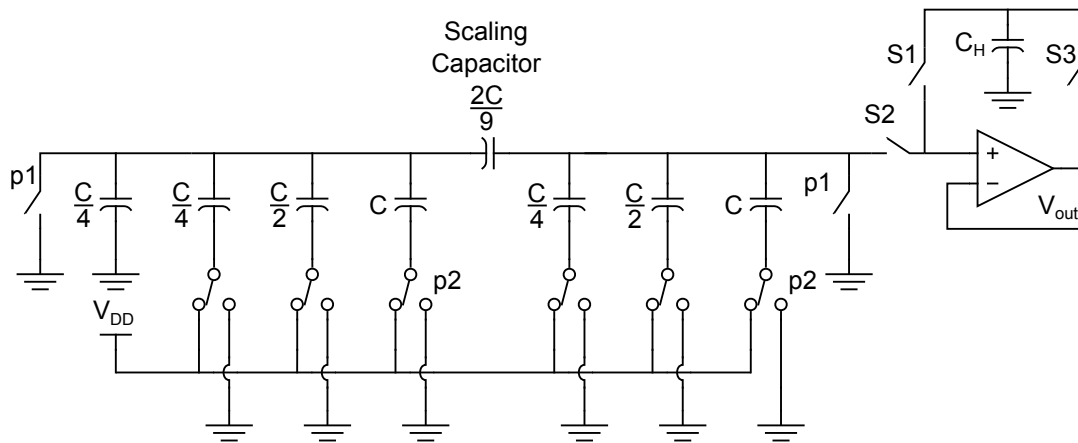
Most other low-power DACs are both higher performance and higher power than desired for our target applications. A 16-bit  $44\ \text{kHz}$  DAC requires  $15\ \text{mW}$  power consumption [107], which is still much too high for our target applications. Other “low-power” approaches have also targeted high speeds [108] [109]. This paucity of research on ultralow-power DACs with relatively low performance requirements is an opportunity for innovation.

## 6.4 DAC Design

The design shown in Figure 6.5 and described in Section 6.3.1 is suitable for 10-bit resolution. Coupled with an opamp that can drive large loads but has relatively low static current draw, this design is almost suitable for our microsystem, but has a serious drawback: the DAC output is only valid during clock phase p2, and returns to ground during phase p1. Because the DAC must output a low-frequency analog waveform to drive the conductivity sensor, the return-to-zero drawback is unacceptable. One way around this problem is to add a sample-and-hold circuit that can capture the output of the DAC during clock phase p2, when the output of the DAC is valid, and then serve as the input to the unity-gain buffer while the other capacitors are discharged. The schematic for this DAC is shown in Figure 6.6.

The design in Figure 6.6 adds three new switches: S1, S2, and S3, to implement a sample-and-hold circuit on the output. These switches must be carefully timed to ensure the DAC operates as planned. It is convenient to trigger these switches on the two-phase clocks p1 and p2. As mentioned in Section 6.3.1, phase p1 is the clock phase during which the switches labeled p1 in Figure 6.6 are closed. Two timing situations need to be considered: the p1-to-p2 transition, and the p2-to-p1 transition.

During clock phase p1, the switches labeled p1 are closed and the DAC capacitors are discharged to ground. The holding capacitor  $C_H$  is connected to the opamp, so S1 is closed, and S2 and S3 are open. To transition to the new output voltage, S1 must open before S2 closes, or the holding capacitor will distort the new output voltage. The inputs to the opamp will be disconnected for a short time, creating a small glitch when transitioning to



**Figure 6.6** Simplified schematic of an extended resolution 6-bit charge scaling DAC with a holding capacitor. The holding capacitor ( $C_H$ ) captures the valid DAC output and preserves that output while the DAC's other capacitors are discharged.

the new output voltage.

During clock phase p2, the switches labeled p1 are open, and S2 is closed to provide the correct analog output voltage to the unity-gain buffer. S1 is open to prevent  $C_H$  from distorting the output voltage, and S3 is closed so that  $C_H$  can capture the output voltage to preserve. Just prior to the transition to phase p1,  $C_H$  will have the same voltage as exists on the other side of S1, so S1 can close before S2 opens. This prevents a glitch from occurring during the transition back to phase p1.

- During p1-p2: S1 open before S2 close
- During p2-p1: S1 close before S2 open

Because the order is the same, S1 and S2 must be triggered by the same clock.

If S3 is triggered by the same clock as S1, then either during the p1-p2 transition S3 will close before S1 opens, or during the p2-p1 transition S1 will close before S3 opens. Either way, the opamp output will be shorted to its input, allowing some drift to occur during that small time window. To avoid this situation, S3 will instead be triggered by a separate clock. The primary constraint in this case is that during the p2 to p1 transition, S3 should open before S2 opens so that the correct output voltage is saved before the opamp input is removed. Therefore, S3 will trigger on p2, while S1 and S2 will trigger on p1.

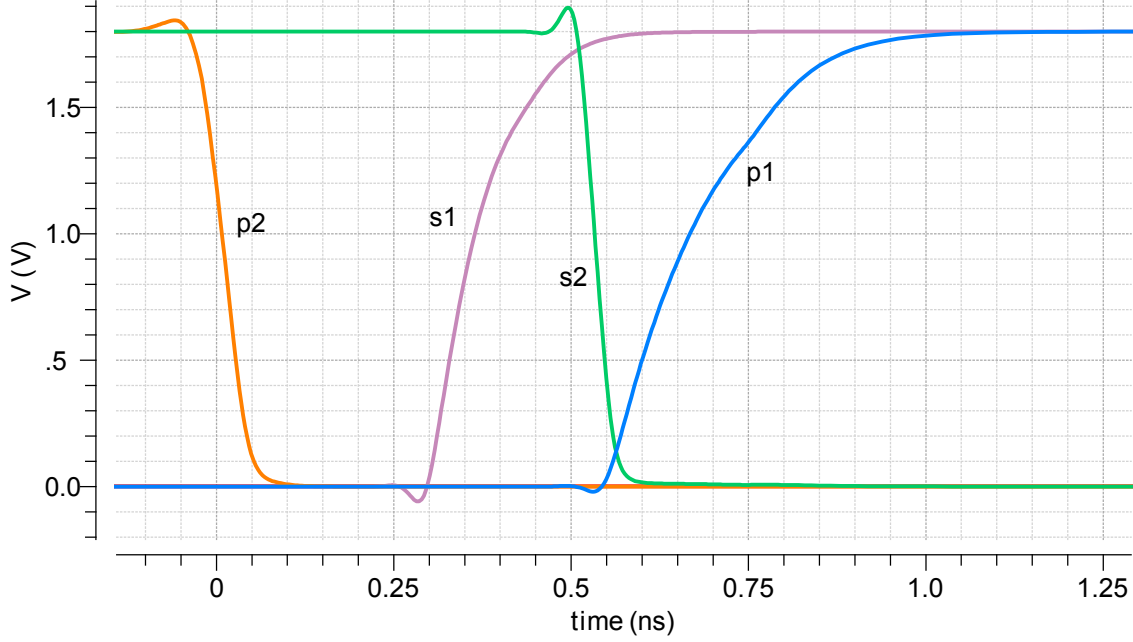
These constraints do not entirely specify timing, but the switch timing used in the final design, summarized in Table 6.2, meets all constraints discussed in this section. A diagram showing the clocks during the p2-p1 transition is shown in Figure 6.7. The diagram for the p1-p2 transition is similar, but the order is s1-s2-p1-p2 and the transitions are to opposite voltages to those in Figure 6.7.

### 6.4.1 Capacitor Sizing

The equation for capacitor thermal noise (or  $kTC$  noise), also used in Section 5.3.3.1 in (5.21), is:

**Table 6.2** DAC switch timing. Two-phase nonoverlapping clocks trigger all switches in the design. This table lists which clock triggers each switch and the order of the switches. An arrow indicates a small delay.

Clock	Switch order
p1	S1 → S2 → p1
p2	p2 and S3 are simultaneous



**Figure 6.7** Timing diagram of DAC's four clocks during p2-p1 transition.

$$v_n^2 = \frac{2kT}{C} \quad (6.7)$$

As in Section 5.3.3.1, the normal equation is doubled because noise can be accumulated from both the charging and discharging events that occur in one clock period. The thermal noise should be compared to the LSB voltage, which is:

$$\text{LSB} = \frac{1.8 \text{ V}}{2^{10}} = 1.76 \text{ mV} \quad (6.8)$$

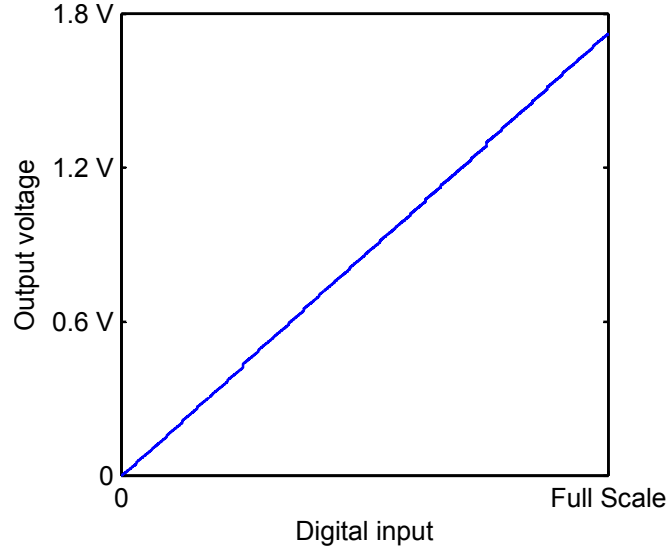
To keep capacitor thermal noise below this value, every capacitor must be larger than

$$C = \frac{2kT}{v_n^2} = 2.67 \text{ fF} \quad (6.9)$$

To substantially exceed this limit, the smallest capacitors are sized at 125 fF. Simulation revealed that larger capacitors achieve more linear output, so capacitors are sized significantly larger than the minimum size calculated above.

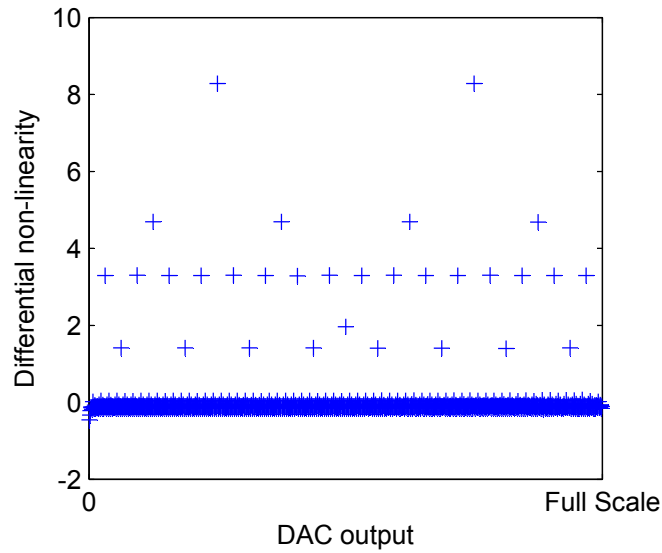
## 6.5 Simulated Performance

The simulated output of the DAC is displayed in Figure 6.8. Although the output is mostly linear, there are some nonlinearities in the output due to parasitic capacitance that exist between the substrate and the vertical natural capacitor. To verify that this parasitic capacitance is the cause of the nonlinearities, simulations with ideal capacitors



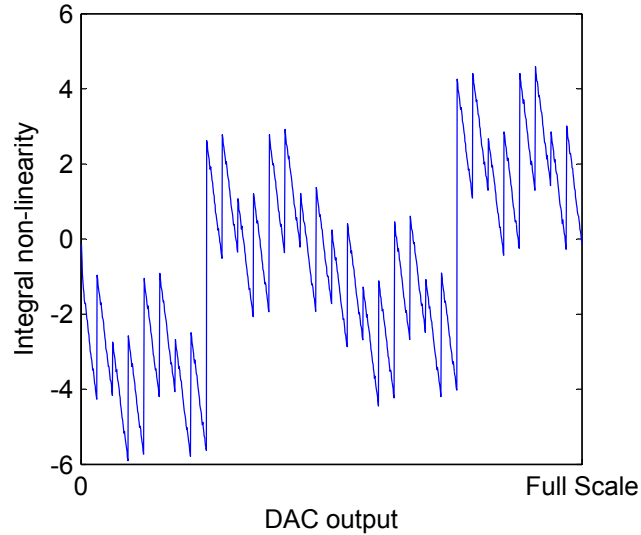
**Figure 6.8** Simulated output of DAC. Displays monotonicity, output range, and gain error.

and parameterizable models of vertical natural capacitors were compared. When simulations are performed with ideal capacitors, the output behaves in accordance with (6.6). However, when the ideal capacitors are replaced with vertical natural capacitor models, the nonlinearities shown in Figures 6.9 and 6.10 result. The effects of the parasitic capacitance could have been corrected by using parasitic-extracted simulations, but were not discovered until after fabrication. The offset of the DAC is only 0.45 mV, which is very near zero, and



**Figure 6.9** Simulated differential non-linearity of DAC. Exposes a small number of nonlinearities due to parasitic capacitance.





**Figure 6.10** Simulated integral nonlinearity of DAC. Exposes a small number of nonlinearities due to parasitic capacitance.

the gain error is 0.04 of the full-scale output. The output range is 0.45 mV to 1.72 V. The simulated DAC output is shown in Figure 6.8.

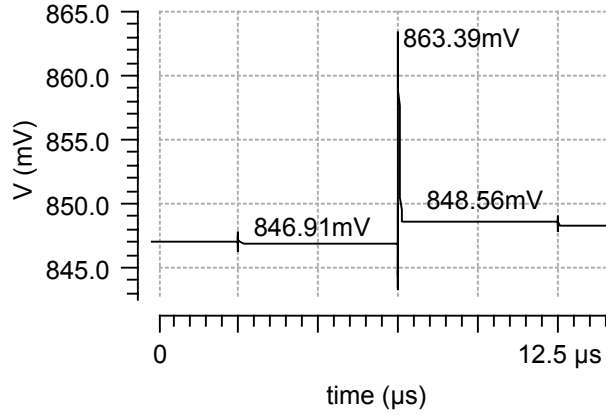
The most important figure of merit for a DAC is the differential nonlinearity (DNL). The DNL is usually reported as a single number, the maximum found for the DAC. The maximum DNL for this DAC is  $8.3 \cdot \text{LSB}$ . Fortunately, the typical transition is much better: 97% of all transitions are  $< 0.5 \cdot \text{LSB}$ , with a small number of transitions with a large nonlinearity when several small capacitors are switched for a single larger capacitor. The DNL across the input range is depicted in Figure 6.9. Parasitic-extracted simulation along with tweaking capacitor sizes could have significantly mitigated this problem. The worst integral nonlinearity (INL) is  $6 \cdot \text{LSB}$ , as shown in Figure 6.10.

Section 6.4 mentioned that a glitch in the output from phase p1 to p2 is inevitable. This glitch is most evident in simulation when the opamp is driving an infinite resistance. This worst-case simulation is depicted in Figure 6.11.

The glitch is less than 17 mV and lasts less than 130 ns. The glitch impulse area (or glitch energy) is therefore 1.1 nV/s. The short duration indicates that its frequency content is very high. When the DAC is driving a 100 pF load through 17 k $\Omega$ , an estimated probe impedance, the glitch entirely disappears.

### 6.5.1 Simulation Result Summary

The simulation results are summarized (alongside the measured results) in Table 6.3. While the DAC had a relatively large DNL and INL, the source of these problems has been



**Figure 6.11** Waveform from DAC simulation showing output glitch. The waveform shows a 17 mV glitch that occurs from p1 to p2. A much smaller glitch occurs from p2 to p1.

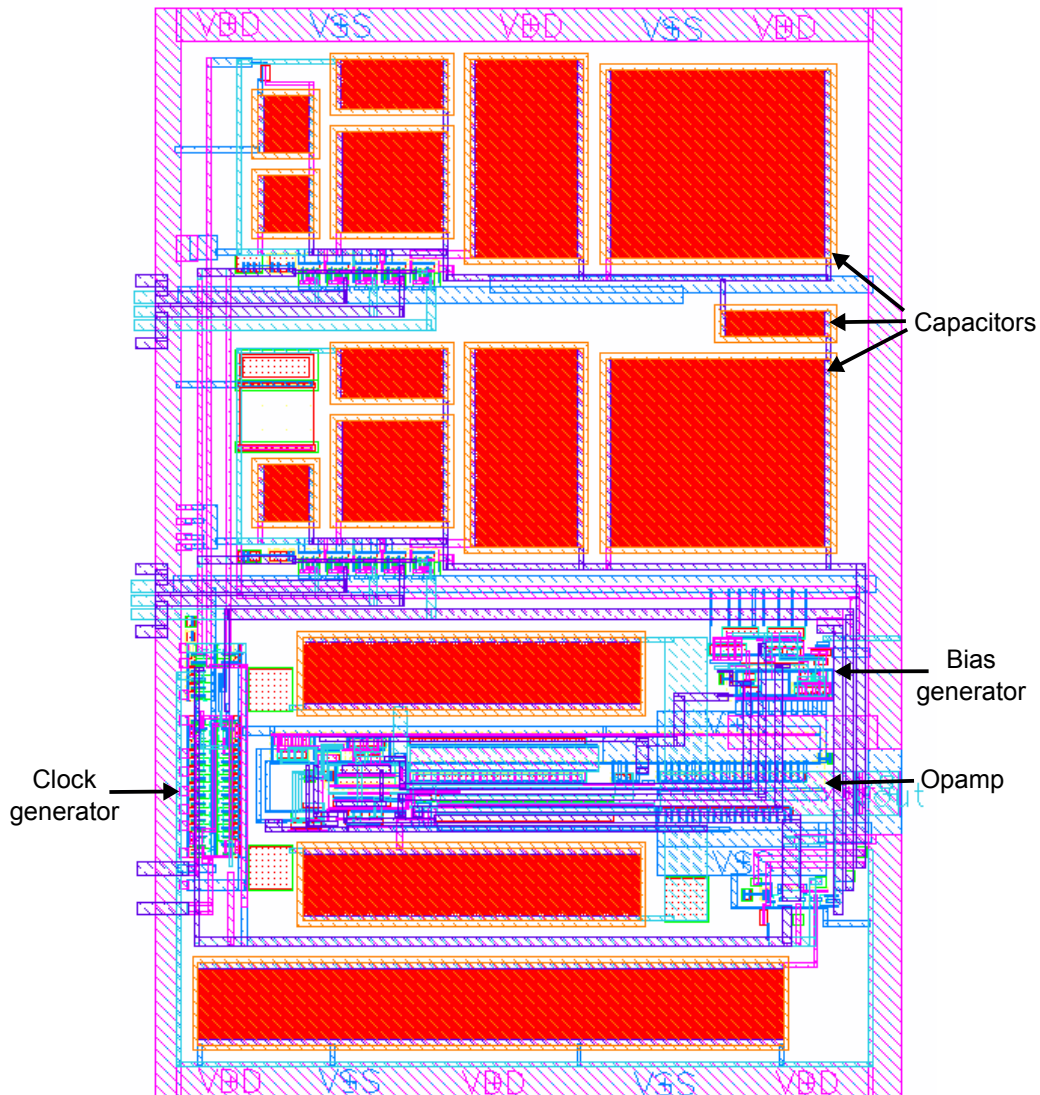
**Table 6.3** DAC simulated and measured results. Power could not be measured because the output was tied to ground through a resistor. The glitch area also could not be measured because of clocking issues described in Section 6.6.

Metric	Simulated	Measured
DNL	$8.3 \cdot \text{LSB}$	$7.3 \cdot \text{LSB}$
INL	$6.0 \cdot \text{LSB}$	$8.0 \cdot \text{LSB}$
Output range	0.45 mV to 1.72 V	4.4 mV to 1.74 V
Monotonic	✓	
Offset	0.45 mV	4.4 mV
Gain error	0.04 of FSO (full-scale output)	0.03 of FSO
Power consumption	109 $\mu\text{W}$ at 1.8 V	8.1 $\mu\text{W}$
Drive current	1.65 mA	1.5 mA
Glitch area	1.1 nV/s	N/A

identified as parasitic capacitance between the substrate and the vertical natural capacitors, and can be mitigated by using parasitic-extracted simulation to adjust capacitor sizes for a future revision. The power consumption is impressively low, and yet high drive current is still achieved. The DAC requires a single opamp, which is reused in other circuits in the microcontroller, and demonstrates a unique design. This design uses a sample-and-hold circuit to keep the output valid across both phases of the clock, while using the sole opamp to buffer the output.

## 6.6 Measured Performance

The layout of the fabricated DAC is shown in Figure 6.12. Measurements of the fabricated DAC revealed two design issues:



**Figure 6.12** Layout of fabricated DAC.

1. There is insufficient buffering between certain clock signals.
2. A diode, intended to be an antenna diode, inadvertently forms a resistor between the DAC output and the substrate.

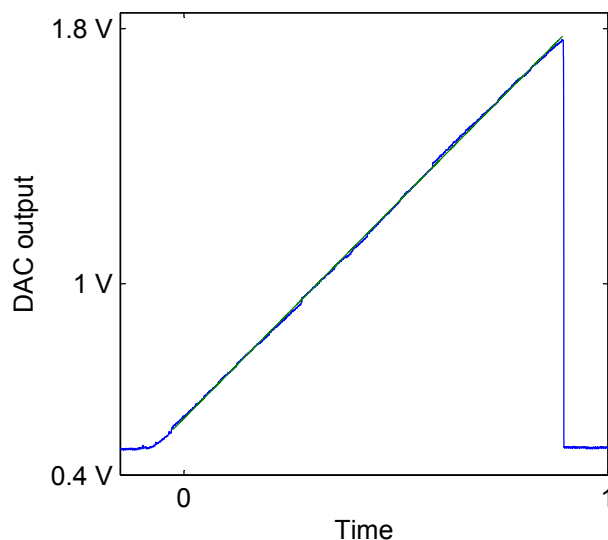
The first issue caused the output to fall towards ground when the input clock was switched. Although the two primary clocks, p1 and p2, are separated by 18 gate delays, some of the derivative clocks, such as s1 and s2, are separated by only 3 gate delays. While this appears to be sufficient in nominal simulation, measured results show that applying the clock signal causes the output to droop towards ground. Recent simulations show that increasing the power of the PMOS transistors by only a single standard deviation creates a

similar effect in simulation. Fortunately, the capacitors retain their charge long enough that when the inputs are switched while keeping the clock input high, the output is reasonably accurate. The DAC is clocked only when all inputs are zeros.

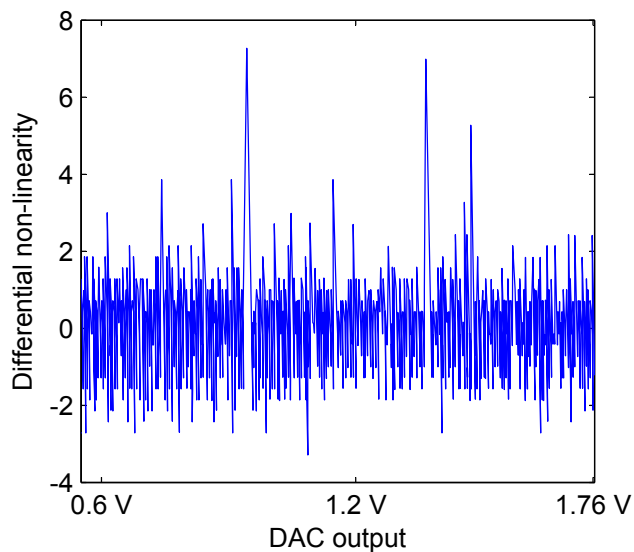
The second issue is that a substrate tie forms a resistor between the DAC output and ground of  $\sim 320\ \Omega$ . The DAC output current is not sufficient to drive the output to 1.8 V, so when a ramp input is applied, the output plateaus when maximum output current is reached. To correct this issue, a 4 mA current source is applied to the output node. With the current source applied, the DAC output cannot drive the output node to ground, but the output reaches 1.74 V.

When each of the above corrective measures is taken, the output shown in Figure 6.13 is achieved. The DNL and INL across the measured output range are depicted in Figures 6.14 and 6.15. The maximum DNL is  $7.3 \cdot \text{LSB}$ , and maximum INL is  $8.0 \cdot \text{LSB}$ . These results match well with simulations, although more noise is present.

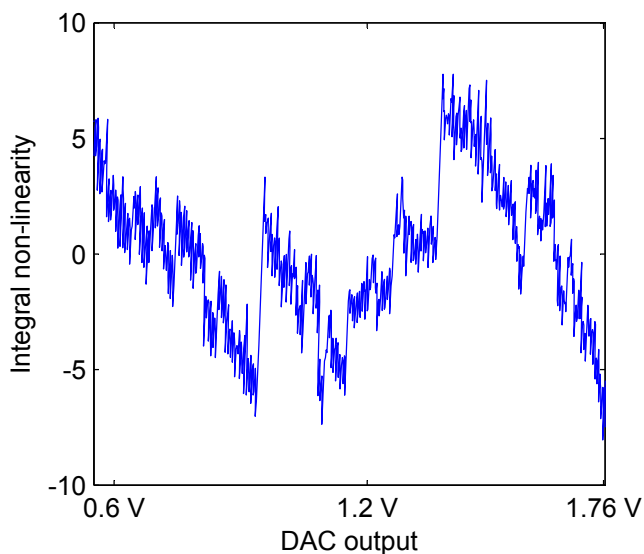
The occurrence of the maximum DNL values at the significant switching voltages of 900 mV and 1.35 V strongly indicates that these values are either due to the parasitic capacitance effects discussed in Section 6.5 or to capacitor mismatch. INL and DNL of this magnitude correspond to a DAC with  $\sim 7$ - to 8-bit accuracy. This accuracy could be increased by correcting for the parasitic capacitance effects discussed in Section 6.5, and using layout techniques to minimize capacitor mismatch.



**Figure 6.13** Section of DAC output with ramp input. Output ranges from .45 V to 1.74 V, with 4 mA current source connected to output. The superimposed green line is the best fit to the measured output.



**Figure 6.14** Measured differential nonlinearity of the DAC output depicted in Figure 6.13.



**Figure 6.15** Measured integral nonlinearity of the DAC output depicted in Figure 6.13.

Many other DNL values may be the result of environmental noise. Direct measurement of the power supply by the oscilloscope showed a periodic signal, with an amplitude of 5 mV peak-to-peak and frequency of 770 Hz. As the LSB voltage is only 1.7 mV, this noise, whether due to the power supply, the oscilloscope, or picked up by unshielded wires, is likely responsible for many of the DNL values in the  $-2$  to  $2 \cdot \text{LSB}$  range. If this environmental noise becomes a limiting factor in a future revision, it can be reduced by using shielded connections to a Faraday cage, using a lower noise power supply, or taking measurements

more slowly to allow the oscilloscope to achieve higher accuracy.

The power of the fabricated DAC was taken with the input set to 0. However, even driving a small offset voltage across a  $320\ \Omega$  output resistor can result in a significant increase in power consumption relative to very low power draw of the opamp. The measured results are summarized with the simulated results in Table 6.3.

## 6.7 Conclusion

All common DAC architectures are considered for the unique requirements of the target applications, and the charge scaling architecture is found to be the most suitable. The simulated performance is hampered by parasitic capacitance, but that problem can be significantly mitigated in a future revision based on parasitic extracted simulation. Several issues affect the fabricated design, but when corrective actions are taken, the DAC achieves  $\sim 7$ - to 8-bit accuracy. The simulated and measured results indicate that the design is feasible for the target applications. Solutions for each issue affecting the design are presented. This design reuses circuits to a high degree, supports a large drive current, and is suitable for a low-power microsystem.

## CHAPTER 7

### CONCLUSION

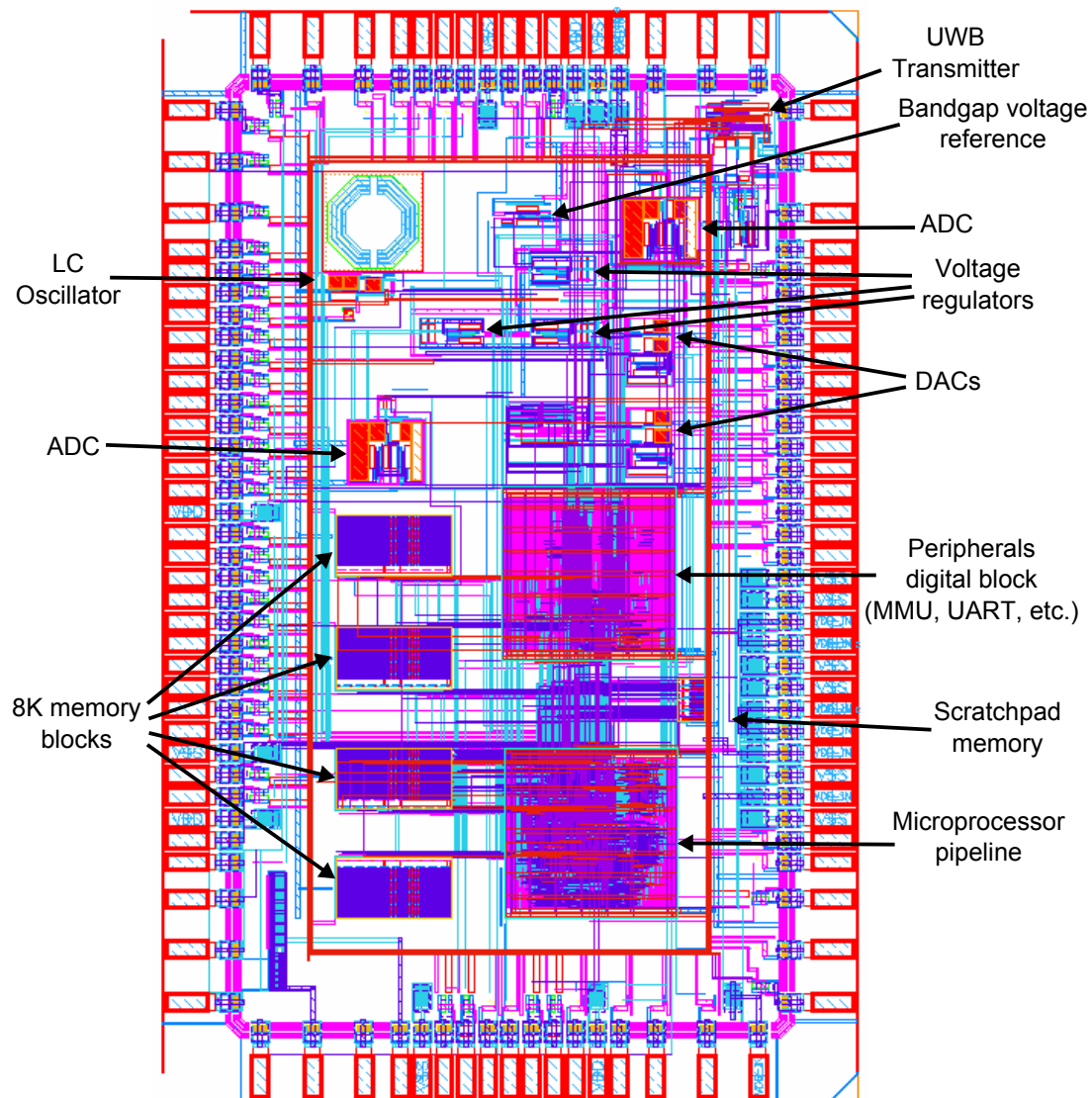
#### 7.1 Motivation

Progress in the field of small low-power microsystems attracts interest because of its enormous potential to increase quality of life. As process scaling enables microchips to become increasingly pervasive, entirely new biomedical applications become possible. In the “More than Moore” era, integration of a large variety of circuitry can add more functionality than is afforded by progress in process technology. The work described in this dissertation is intended to demonstrate that motivated researchers can integrate all the circuitry necessary for many of these emerging biomedical applications into a single chip.

To give concrete objectives to these lofty goals, two specific applications were chosen: an S-IVR, and an NO-releasing catheter. Both of these applications have the potential to make significant medical contributions. The S-IVR would enable the testing of an AIDS protection device that could help slow the spread of AIDS. The NO-releasing catheter could help reduce many of the thousands of infections associated with catheters. These applications provide specific requirements for a prototype biomedical microsystem, yet represent a broad class of emerging biomedical applications. These two applications share the requirements of small physical size, long battery life, sensor interfaces, and a wireless transceiver. The small physical size and long battery life in turn necessitate single-chip integration and low-power circuit design. The sensor interfaces can support a variety of micropotentiometric, resistive (DC), and conductive (AC) sensors.

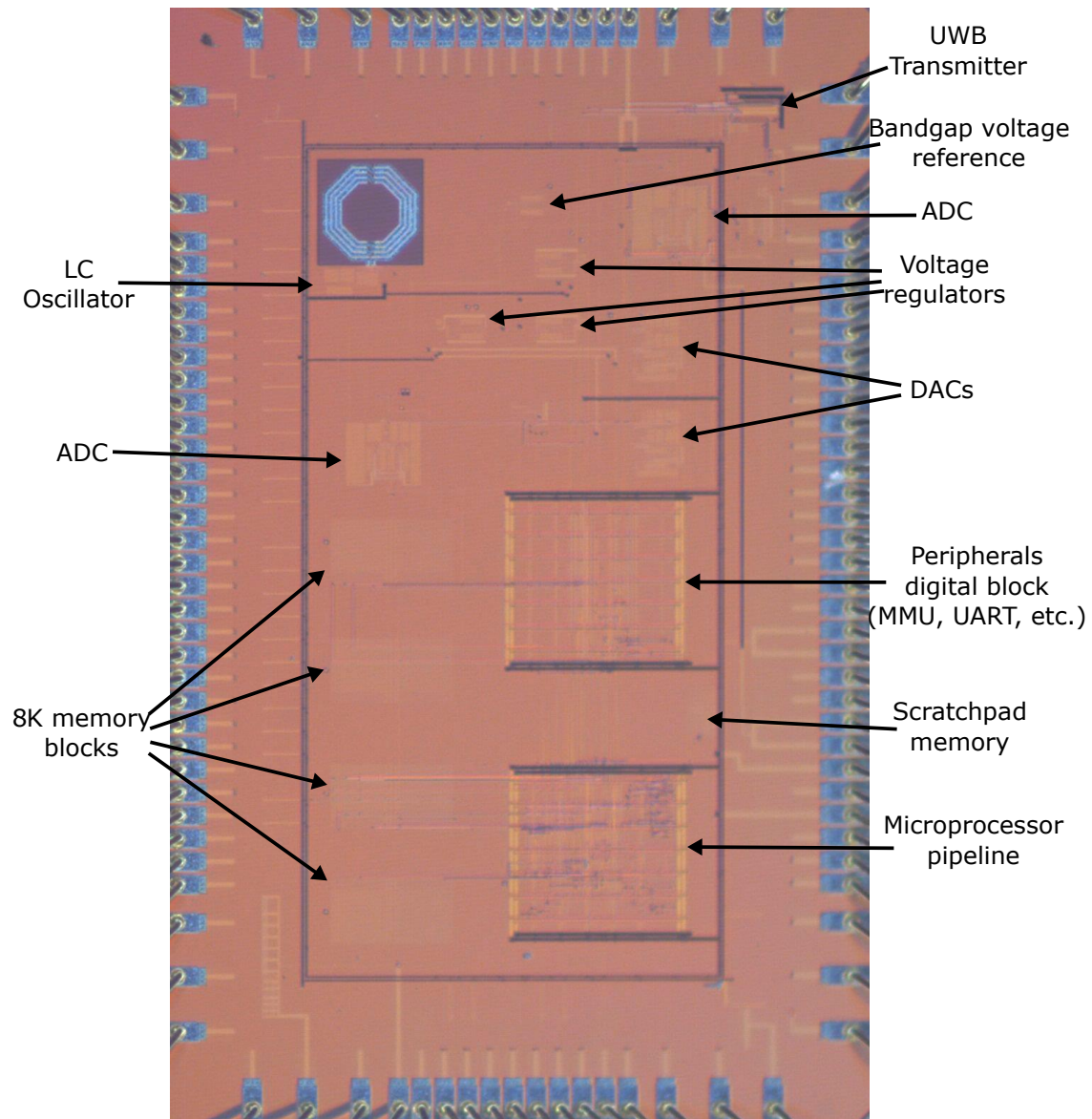
#### 7.2 Dissertation Accomplishments

The crowning achievement of the work described here is the definition, design, fabrication and testing of the Utah Microcontroller, which integrates all of the necessary circuitry for both of the targeted applications. The layout of the Utah Microcontroller is depicted in Figure 7.1, and a die photo in Figure 7.2. This is the latest in a family of microcontrollers developed in our research group, and is built upon the architectural, software and testing



**Figure 7.1** Layout of the Utah Microcontroller. Many of the most significant blocks are identified.





**Figure 7.2** Die photo of the Utah Microcontroller. Many of the most significant blocks are identified.

tools of previous students. As mentioned in the body of the dissertation, the Utah Microcontroller was developed in close collaboration with Ondrej Novak. He contributed designs for voltage regulators, circuitry to enable a microcontroller sleep mode, an inductive link to wirelessly recharge the microsystem battery, and a UWB (ultrawideband) transmitter [17].

The UWB transmitter is all-digital, which results in very low standby power, instantaneous startup, and low power consumption. UWB provides large bandwidth and large maximum data rates, and reduces power consumption at the expense of increased complexity at the receiver. This is ideal for biomedical applications, where the receiver is not resource constrained [17].

The Utah Microcontroller was implemented in a 65 nm technology (earlier versions had been in 180 nm processes), facilitating the analysis of architectural features with scaling. A test flow was developed to enable testing on a Verigy 93000 and accurate power dissipation measurements. A novel method to measure per-instruction energy [39] was used to characterize new microcontroller instructions. The work in Section 2.3 analyzed the costs of supporting a 2 MB address space instead of a 64 kB address space. That research concluded that the address space expansion incurred an area penalty of only 2.5%, a leakage penalty of 3%, and a similarly small increase in active energy.

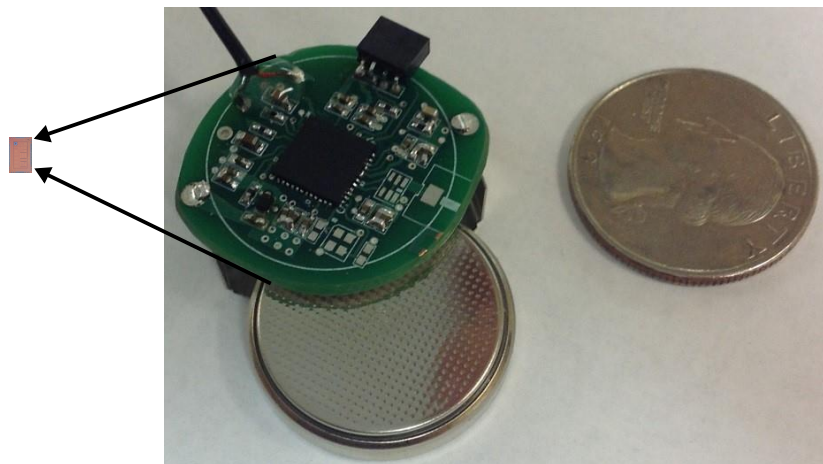
The work described in Chapter 3 considered how the WIMS Microprocessor's scratchpad memory scaled from 180 nm to 65 nm. It also considered broader scaling trends, and projected the benefits of the scratchpad memory into future process nodes. In a 32 nm version of the WIMS Microprocessor, the scratchpad memory is projected to save  $\sim 10$ –30% of memory access energy depending upon the characteristics of the embedded program.

This dissertation demonstrates that it is possible to integrate the analog circuitry needed to interface to a wide variety of solid-state sensors. The specific sensors targeted by this dissertation are a pH sensor, a temperature sensor, and a conductivity sensor. However, the interfaces required by each of these sensors can be applied to a broad variety of sensors required by similar small low-power sensing applications. The pH sensor is a micropotentiometric sensor. Many ion-selective sensors for detecting concentrations of different ions also require the same high input impedance interface. The interface for the temperature sensor supports any sensor that modulates a resistance, or converts a constant current source into a voltage. The interface for the conductivity sensor can likewise support any sensor that modulates impedance in the same frequency range. These three sensor interfaces collectively support a broad variety of chemical and biological sensors, and the Utah Microcontroller could be used for many applications beyond those specifically targeted

in the dissertation.

The analog portion of the microcontroller is a demonstration of what can be accomplished in a design style that is focused on circuit reuse. For example, the opamp described in Chapter 4 is used in the integrators in the sigma-delta ADC, the high input impedance buffer required for the pH sensor, and the high output current buffer in the DAC. Although the fabricated opamp does not achieve the anticipated bandwidth or slew rate, the design error and layout effects that result in the reduction of these metrics have been identified. The opamp does achieve very low power consumption, high gain, and high output drive strength. The DAC illustrates a unique (to the knowledge of the author) topology, using a sample-and-hold circuit to form a single-opamp charge-scaling DAC.

Figure 7.3 compares the size of a PCB-based microsystem and the Utah Microcontroller. The PCB-based microsystem was developed to power an NO-releasing catheter, while the Utah Microcontroller can be used for that or many other small sensing microsystems. The reduction in area is over  $300\times$ , and a reduction in volume of over  $1000\times$ . Integration also reduces overall power consumption by eliminating I/O power between the integrated components. The decreased power and volume enable emerging biomedical applications that can increase quality of life for thousands to millions of people. Several examples of these emerging applications are given in Chapter 1. Many more commercial applications involving wearable computing and the “Internet of Things” will arise if wireless capability, useful sensors, energy harvesting, and very small physical size can be combined.



**Figure 7.3** Size comparison of a PCB-based microsystem and the Utah Microcontroller. Both microsystems are shown approximately at scale.

### 7.3 Future Work

Future work can be broken down into two different potential objectives:

1. Realizing a prototype microsystem for a particular biomedical application.
2. Enabling the development of the types of implantable biomedical microsystems targeted in this dissertation.

The work in this dissertation has focused on the development of a chip capable of supporting a variety of biomedical applications, but the emerging biomedical applications targeted in this dissertation require further work to be realized. Another iteration of the Utah Microcontroller would enable the design fixes identified in Chapters 4, 5, and 6, which would substantially improve the analog circuits described in those chapters.

The core component of the Utah Microcontroller is the WIMS Microprocessor, which is examined in Chapters 2 and 3. This microprocessor has been improved by several generations of graduate students, and has reached a point of high reliability with a strong supporting toolset. Pad issues have prevented the thorough testing of new and improved digital components, such as an improved UART. This UART enables in-the-field programming of the microcontroller, and remains a high priority for future development.

Wireless programming of the microcontroller would also add significant advantages. For small systems such as an S-IVR, without a wireless programming option, programming must take place before the manufacture of the S-IVR. With wireless programming, along with an inductive link for wireless charging, manufacturing can take place once, in bulk, and each S-IVR could be reprogrammed as needed by researchers. This would also enable quick sensor calibration updates over the course of a study. In case the battery completely depletes, the wireless receiver could be able to apply a hard reset and bring the microcontroller into a known state. Wireless programming would benefit many implantable biomedical applications in which packaging prevents convenient physical access to a serial port.

There are also options for further lowering power consumption. If the microprocessor was implemented with a custom digital library minimizing the number of transistors stacked between  $V_{DD}$  and ground, the microprocessor could be run at even lower voltages and consume less power. A scratchpad-based low-power mode could set the main memory banks into a low-voltage and low-power retention state and execute only from the scratchpad memory. This mode would also set the voltage regulators to decrease the supply voltage of the processor, and decrease the frequency of the on-chip clock.

Energy harvesting could further extend the battery life of implantable biomedical microsystems. Energy harvesting techniques usually output very different voltage levels (usually AC) than supported by the voltage regulators in the Utah Microcontroller (designed for DC battery inputs), so custom power conversion circuits are required to incorporate energy harvesting [110]. These techniques could further enhance the suitability of the Utah Microcontroller for many low power applications.

Steps have also been made toward the second potential objective, to enable the development of small low-power microsystems. For example, the WIMS Microprocessor is freely available on UMIPS (University of Michigan Intellectual Property Source), and may benefit researchers pursuing the development of highly integrated low-power microsystems. The analog components in the Utah Microcontroller are well documented in Chapters 4, 5, and 6. The circuit schematics and explanations of the principles of operation given in these chapters will be of use to anyone seeking to reproduce these important circuits. Many further steps can also be taken. The test environments and MATLAB code used to simulate and characterize the DAC and sigma-delta ADC could be released under an open source license.

As analog circuit design matures, designs are becoming increasingly standardized. This development is a boon to microsystem designers, as the primary challenges now lie in integrating these circuits into application-driven microsystems. This presents the opportunity of creating open source versions of common designs in a freely available process technology, such as the Predictive Technology Model [74], which simulates advanced process nodes. Such open source designs, accompanied by open source test environments and analysis code, could significantly accelerate the development of suitable analog circuits by microsystem researchers.

## 7.4 Conclusion

This work has described a highly integrated mixed-signal microsystem. To give concreteness to the design goals, two specific microsystems were targeted: an NO-releasing catheter and an S-IVR, as described in Chapter 1. These two target applications created a set of requirements that are representative of a broad class of biomedical sensor systems. These requirements led to a system that sets a new level of integrated functionality for single-chip microcontrollers. The Utah Microcontroller contains low-power interfaces for a broad variety of sensors, a high-efficiency wireless transceiver, sleep mode, and sufficient processing power for on-chip data analysis and compression. This dissertation is an example

of how to survey a large variety of circuit architectures, select the circuit best suited to the desired specifications, and optimize the circuit to those specifications. Each of the necessary circuits has been designed and characterized. Finally, all circuitry necessary to support these applications has been integrated and fabricated on a single chip: the Utah Microcontroller.

## APPENDIX

### CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL PUBLIC LICENSE

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International Public License (“Public License”). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

#### Section 1 – Definitions.

- a. **Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.
- b. **Adapter’s License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.
- c. **Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)–(2) are not Copyright and Similar Rights.

- d. **Effective Technological Measures** means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.
- e. **Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.
- f. **Licensed Material** means the artistic or literary work, database, or other material to which the Licensor applied this Public License.
- g. **Licensed Rights** means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.
- h. **Licensor** means the individual(s) or entity(ies) granting rights under this Public License.
- i. **Share** means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.
- j. **Sui Generis Database Rights** means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.
- k. **You** means the individual or entity exercising the Licensed Rights under this Public License. Your has a corresponding meaning.

## Section 2 – Scope.

### a. **License grant.**

- 1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive,



irrevocable license to exercise the Licensed Rights in the Licensed Material to:

- A. reproduce and Share the Licensed Material, in whole or in part; and
  - B. produce, reproduce, and Share Adapted Material.
2. Exceptions and Limitations. For the avoidance of doubt, where Exceptions and Limitations apply to Your use, this Public License does not apply, and You do not need to comply with its terms and conditions.
  3. Term. The term of this Public License is specified in Section 6(a).
  4. Media and formats; technical modifications allowed. The Licensor authorizes You to exercise the Licensed Rights in all media and formats whether now known or hereafter created, and to make technical modifications necessary to do so. The Licensor waives and/or agrees not to assert any right or authority to forbid You from making technical modifications necessary to exercise the Licensed Rights, including technical modifications necessary to circumvent Effective Technological Measures. For purposes of this Public License, simply making modifications authorized by this Section 2(a)(4) never produces Adapted Material.
  5. Downstream recipients.
    - A. Offer from the Licensor - Licensed Material. Every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights under the terms and conditions of this Public License.
    - B. No downstream restrictions. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, the Licensed Material if doing so restricts exercise of the Licensed Rights by any recipient of the Licensed Material.
  6. No endorsement. Nothing in this Public License constitutes or may be construed as permission to assert or imply that You are, or that Your use of the Licensed Material is, connected with, or sponsored, endorsed, or granted official status by, the Licensor or others designated to receive attribution as provided in Section 3(a)(1)(A)(i).

**b. Other rights.**

1. Moral rights, such as the right of integrity, are not licensed under this Public License, nor are publicity, privacy, and/or other similar personality rights; however, to the extent possible, the Licensor waives and/or agrees not to assert any such rights held by the Licensor to the limited extent necessary to allow You to exercise the Licensed Rights, but not otherwise.
2. Patent and trademark rights are not licensed under this Public License.
3. To the extent possible, the Licensor waives any right to collect royalties from You for the exercise of the Licensed Rights, whether directly or through a collecting society under any voluntary or waivable statutory or compulsory licensing scheme. In all other cases the Licensor expressly reserves any right to collect such royalties.

### **Section 3 – License Conditions.**

Your exercise of the Licensed Rights is expressly made subject to the following conditions.

#### **a. Attribution.**

1. If You Share the Licensed Material (including in modified form), You must:
  - A. retain the following if it is supplied by the Licensor with the Licensed Material:
    - i. identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);
    - ii. a copyright notice;
    - iii. a notice that refers to this Public License;
    - iv. a notice that refers to the disclaimer of warranties;
    - v. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;
  - B. indicate if You modified the Licensed Material and retain an indication of any previous modifications; and
  - C. indicate the Licensed Material is licensed under this Public License, and include the text of, or the URI or hyperlink to, this Public License.
2. You may satisfy the conditions in Section 3(a)(1) in any reasonable manner based on the medium, means, and context in which You Share the Licensed

Material. For example, it may be reasonable to satisfy the conditions by providing a URI or hyperlink to a resource that includes the required information.

3. If requested by the Licensor, You must remove any of the information required by Section 3(a)(1)(A) to the extent reasonably practicable.
4. If You Share Adapted Material You produce, the Adapter's License You apply must not prevent recipients of the Adapted Material from complying with this Public License.

#### **Section 4 – Sui Generis Database Rights.**

Where the Licensed Rights include Sui Generis Database Rights that apply to Your use of the Licensed Material:

- a. for the avoidance of doubt, Section 2(a)(1) grants You the right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database;
- b. if You include all or a substantial portion of the database contents in a database in which You have Sui Generis Database Rights, then the database in which You have Sui Generis Database Rights (but not its individual contents) is Adapted Material; and
- c. You must comply with the conditions in Section 3(a) if You Share all or a substantial portion of the contents of the database.

For the avoidance of doubt, this Section 4 supplements and does not replace Your obligations under this Public License where the Licensed Rights include other Copyright and Similar Rights.

#### **Section 5 – Disclaimer of Warranties and Limitation of Liability.**

- a. Unless otherwise separately undertaken by the Licensor, to the extent possible, the Licensor offers the Licensed Material as-is and as-available, and makes no representations or warranties of any kind concerning the Licensed Material, whether express, implied, statutory, or other. This includes, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of latent or other defects, accuracy, or the presence or absence of errors, whether or not known or discoverable. Where disclaimers

of warranties are not allowed in full or in part, this disclaimer may not apply to You.

- b. To the extent possible, in no event will the Licensor be liable to You on any legal theory (including, without limitation, negligence) or otherwise for any direct, special, indirect, incidental, consequential, punitive, exemplary, or other losses, costs, expenses, or damages arising out of this Public License or use of the Licensed Material, even if the Licensor has been advised of the possibility of such losses, costs, expenses, or damages. Where a limitation of liability is not allowed in full or in part, this limitation may not apply to You.
- c. The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.

## **Section 6 – Term and Termination.**

- a. This Public License applies for the term of the Copyright and Similar Rights licensed here. However, if You fail to comply with this Public License, then Your rights under this Public License terminate automatically.
- b. Where Your right to use the Licensed Material has terminated under Section 6(a), it reinstates:
  - 1. automatically as of the date the violation is cured, provided it is cured within 30 days of Your discovery of the violation; or
  - 2. upon express reinstatement by the Licensor.

For the avoidance of doubt, this Section 6(b) does not affect any right the Licensor may have to seek remedies for Your violations of this Public License.

- c. For the avoidance of doubt, the Licensor may also offer the Licensed Material under separate terms or conditions or stop distributing the Licensed Material at any time; however, doing so will not terminate this Public License.
- d. Sections 1, 5, 6, 7, and 8 survive termination of this Public License.

## **Section 7 – Other Terms and Conditions.**

- a. The Licensor shall not be bound by any additional or different terms or conditions communicated by You unless expressly agreed.

- b. Any arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.

## **Section 8 – Interpretation.**

- a. For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully be made without permission under this Public License.
- b. To the extent possible, if any provision of this Public License is deemed unenforceable, it shall be automatically reformed to the minimum extent necessary to make it enforceable. If the provision cannot be reformed, it shall be severed from this Public License without affecting the enforceability of the remaining terms and conditions.
- c. No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.
- d. Nothing in this Public License constitutes or may be interpreted as a limitation upon, or waiver of, any privileges and immunities that apply to the Licensor or You, including from the legal processes of any jurisdiction or authority.

## REFERENCES

- [1] G. Chen, H. Ghaed, R. Haque, M. Wieckowski, K. Yejoong, K. Gyouho, D. Fick, K. Daeyeon, S. Mingoo, K. Wise, D. Blaauw, and D. Sylvester, "A cubic-millimeter energy-autonomous wireless intraocular pressure monitor," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), IEEE International*, 2011, Conference Proceedings, pp. 310–312.
- [2] J. Ko, C. Lu, M. B. Srivastava, J. A. Stankovic, A. Terzis, and M. Welsh, "Wireless sensor networks for healthcare," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1947–1960, 2010. [Online]. Available: <http://10.1109/JPROC.2010.2065210>
- [3] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore, "Environmental wireless sensor networks," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1903–1917, 2010. [Online]. Available: <http://10.1109/JPROC.2010.2068530>
- [4] G. Chen, M. Fojtik, K. Daeyeon, D. Fick, P. Junsun, S. Mingoo, C. Mao-Ter, F. Zhiyoong, D. Sylvester, and D. Blaauw, "Millimeter-scale nearly perpetual sensor system with stacked battery and solar cells," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), IEEE International*, 2010, Conference Proceedings, pp. 288–289. [Online]. Available: <http://10.1109/ISSCC.2010.5433921>
- [5] J. Pandolfino, J. Richter, T. Ours, J. Guardino, J. Chapman, and P. Kahrilas, "Ambulatory esophageal pH monitoring using a wireless system," *The American Journal of Gastroenterology*, vol. 98, no. 4, pp. 740–749, 2003.
- [6] J. Pritchett, M. Aslam, J. Slaughter, R. Ness, C. Garrett, and M. Vaezi, "Efficacy of esophageal impedance/pH monitoring in patients with refractory gastroesophageal reflux disease, on and off therapy," *Clinical Gastroenterology and Hepatology*, vol. 7, no. 7, p. 743, 2009.
- [7] R. Tutuian and D. Castell, "Gastroesophageal reflux monitoring: pH and impedance," *GI Motility online*, 2006. [Online]. Available: <http://www.nature.com/gimo/contents/pt1/full/gimo31.html>
- [8] P. A. Hammond, D. Ali, and D. R. S. Cumming, "A system-on-chip digital pH meter for use in a wireless diagnostic capsule," *Biomedical Engineering, IEEE Transactions on*, vol. 52, no. 4, pp. 687–694, 2005.
- [9] "UNAIDS Report on the global AIDS epidemic 2013," p. 82, 2013. [Online]. Available: <http://www.unaids.org/en/resources/campaigns/globalreport2013/globalreport/>
- [10] "Kiser Lab at Northwestern University," 2015. [Online]. Available: <http://faculty.mccormick.northwestern.edu/patrick-kiser/>

- [11] “UNAIDS Report on the global AIDS epidemic 2010,” p. 83, 2010. [Online]. Available: [http://www.unaids.org/globalreport/Global\\_report.htm](http://www.unaids.org/globalreport/Global_report.htm)
- [12] P. Kiser, T. Johnson, and J. Clark, “State of the art in intravaginal ring technology for topical prophylaxis of HIV infection,” *AIDS Reviews*, vol. 14, pp. 62–77, 2012.
- [13] E. E. Tolley, P. F. Harrison, E. Goetghebeur, K. Morrow, R. Pool, D. Taylor, S. N. Tillman, and A. van der Straten, “Adherence and its measurement in phase 2/3 microbicide trials,” *AIDS and Behavior*, vol. 14, no. 5, pp. 1124–36, 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19924525>
- [14] L. Osterberg and T. Blaschke, “Adherence to medication,” *New England Journal of Medicine*, vol. 353, no. 5, pp. 487–497, 2005. [Online]. Available: <http://dx.doi.org/10.1056/NEJMr050100>
- [15] P. Řezáč, “Potential applications of electrical impedance techniques in female mammalian reproduction,” *Theriogenology*, vol. 70, no. 1, pp. 1–14, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0093691X08001088>
- [16] J. Rodrigues, J. Caldeira, and B. Vaidya, “A novel intra-body sensor for vaginal temperature monitoring,” *Sensors*, vol. 9, no. 4, pp. 2797–2808, 2009. [Online]. Available: <http://www.mdpi.com/1424-8220/9/4/2797/>
- [17] O. Novak, “Radio frequency-power mixed-signal microcontroller with wireless ultra-wideband transmitter for electrochemistry and biosensing,” Ph.D. dissertation, University of Utah, 2014.
- [18] L. A. Mermel, B. M. Farr, R. J. Sherertz, I. I. Raad, N. O’Grady, J. S. Harris, and D. E. Craven, “Guidelines for the management of intravascular catheter-related infections,” *Clinical Infectious Diseases*, vol. 32, no. 9, pp. 1249–1272, 2001. [Online]. Available: <http://cid.oxfordjournals.org/content/32/9/1249.short>
- [19] I. Raad, W. Costerton, U. Sabharwal, M. Sadlowski, E. Anaissie, and G. P. Bodey, “Ultrastructural analysis of indwelling vascular catheters: a quantitative relationship between luminal colonization and duration of placement,” *Journal of Infectious Diseases*, vol. 168, no. 2, pp. 400–407, 1993.
- [20] R. M. Donlan, “Biofilms and device-associated infections,” *Emerging Infectious Diseases*, vol. 7, no. 2, p. 277, 2001.
- [21] H. Zhang, G. M. Annich, J. Miskulin, K. Stankiewicz, K. Osterholzer, S. I. Merz, R. H. Bartlett, and M. E. Meyerhoff, “Nitric oxide-releasing fumed silica particles: synthesis, characterization, and biomedical application,” *Journal of the American Chemical Society*, vol. 125, no. 17, pp. 5015–5024, 2003, pMID: 12708851. [Online]. Available: <http://pubs.acs.org/doi/abs/10.1021/ja0291538>
- [22] L. Höfler, D. Koley, J. Wu, C. Xi, and M. E. Meyerhoff, “Electromodulated release of nitric oxide through polymer material from reservoir of inorganic nitrite salt,” *RSC Adv.*, vol. 2, pp. 6765–6767, 2012. [Online]. Available: <http://dx.doi.org/10.1039/C2RA20853A>
- [23] M. W. Radomski, R. M. Palmer, and S. Moncada, “An l-arginine/nitric oxide pathway present in human platelets regulates aggregation,” *Proceedings of the National Academy of Sciences*, vol. 87, no. 13, pp. 5193–5197, 1990. [Online]. Available: <http://www.pnas.org/content/87/13/5193.abstract>

- [24] L. J. Ignarro, G. M. Buga, K. S. Wood, R. E. Byrns, and G. Chaudhuri, "Endothelium-derived relaxing factor produced and released from artery and vein is nitric oxide," *Proceedings of the National Academy of Sciences*, vol. 84, no. 24, pp. 9265–9269, 1987. [Online]. Available: <http://www.pnas.org/content/84/24/9265.abstract>
- [25] H. Ren, A. Colletta, D. Koley, J. Wu, C. Xi, T. C. Major, R. H. Bartlett, and M. E. Meyerhoff, "Thromboresistant/anti-biofilm catheters via electrochemically modulated nitric oxide release," *Bioelectrochemistry*, vol. 104, no. 0, pp. 10 – 16, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1567539414001959>
- [26] O. Novak, B. Redd, and R. B. Brown, "A fully integrated power-efficient SoC with a wireless uwb transmitter for biomedical and chemical research," in *Proc. International SoC Design Conference (ISOCC)*, 11 2013, Conference Proceedings, pp. 107–110.
- [27] "Silicon labs parametric search," March 2014. [Online]. Available: <http://www.silabs.com/support/Pages/ParametricSearch.aspx?searchType=MCUs>
- [28] *Wireless MCU Solutions*, Silicon Labs, 2012.
- [29] "Silicon Labs Wireless MCUs," March 2014. [Online]. Available: <http://www.silabs.com/products/wireless/wirelessmcu/Pages/default.aspx>
- [30] *rfPIC12F675*, Microchip Technology Inc., 2003–2013.
- [31] *Si106x/108x*, 1st ed., Silicon Labs, February 2014.
- [32] *NFC ISO15693 Sensor Transponder*, Texas Instruments Inc., 2012.
- [33] *CC430 Family User's Guide (SLAU259E)*, Texas Instruments Inc., January 2013.
- [34] *MSP430<sup>TM</sup> SoC With RF Core (SLAS554H)*, Texas Instruments Inc., September 2013.
- [35] *CC1110Fx / CC1111Fx (SWRS033H)*, Texas Instruments Inc., February 2013.
- [36] B. Redd, S. Kellis, N. Gaskin, M. Guthaus, and R. Brown, "Architecture for increased address space in an ultra-low-power microprocessor," in *Circuits and Systems (MWSCAS), 2012 IEEE 55th International Midwest Symposium on*, Aug 2012, pp. 125–128.
- [37] R. M. Senger, E. D. Marsman, M. S. McCorquodale, F. H. Gebara, K. L. Kraver, M. R. Guthaus, and R. B. Brown, "A 16-bit mixed-signal microsystem with integrated CMOS-MEMS clock reference," pp. 520–525, 2003.
- [38] E. D. Marsman, R. M. Senger, G. A. Carichner, S. Kubba, M. S. McCorquodale, and R. B. Brown, "DSP architecture for cochlear implants," in *Circuits and Systems, IEEE International Symposium on*, 2006, Conference Proceedings, p. 4. [Online]. Available: <http://web.eecs.umich.edu/~mmccorq/publications/marsmanISCAS06.pdf>
- [39] S. Kellis, N. Gaskin, B. Redd, J. Campbell, and R. Brown, "Energy profile of a microcontroller for neural prosthetic application," in *Circuits and Systems, IEEE International Symposium on*, May 2010, Conference Paper, pp. 3841–3844.
- [40] (2015) MOSIS Accounts for Academic Institutions. [Online]. Available: <https://www.mosis.com/you-are/academic-institutions>



- [41] R. A. Ravindran, P. D. Nagarkar, G. S. Dasika, E. D. Marsman, R. M. Senger, S. A. Mahlke, and R. B. Brown, "Compiler managed dynamic instruction placement in a low-power code cache," in *Proc. Int. Symp. Code Generation Optimization (CGO)*, 2005, Conference Proceedings, pp. 179–190.
- [42] B. Redd, S. Kellis, N. Gaskin, and R. Brown, "Scratchpad memories in the context of process scaling," in *IEEE 54th Int. Midwest Symp. on Circuits and Systems (MWSCAS)*, Aug 2011, Conference Proceedings.
- [43] R. A. Ravindran, R. M. Senger, E. D. Marsman, G. S. Dasika, M. R. Guthaus, S. A. Mahlke, and R. B. Brown, "Partitioning variables across register windows to reduce spill code in a low-power processor," *Computers, IEEE Transactions on*, vol. 54, no. 8, pp. 998–1012, 2005.
- [44] R. Fleck, B. Holmer, R. Arnold, and B. Singh, "Processing unit having independent execution units for parallel execution of instructions of different category with instructions having specific bits indicating instruction size and category respectively," US Patent US 6 292 845 B1, September 18, 2001. [Online]. Available: <https://www.google.com/patents/US6292845>
- [45] D. Brooks and M. Martonosi, "Dynamically exploiting narrow width operands to improve processor power and performance," in *High-Performance Computer Architecture, Fifth International Symposium On*, January 1999, Conference Proceedings, pp. 13–22.
- [46] B. Redd, S. Kellis, N. Gaskin, and R. Brown, "The impact of process scaling on scratchpad memory energy savings," *Journal of Low Power Electronics and Applications*, vol. 4, no. 3, pp. 231–251, 2014. [Online]. Available: <http://www.mdpi.com/2079-9268/4/3/231>
- [47] S. Steinke, L. Wehmeyer, L. Bo-Sik, and P. Marwedel, "Assigning program and data objects to scratchpad for energy reduction," in *Proc. DATE*, 2002, Conference Proceedings, pp. 409–415.
- [48] J. Kin, M. Gupta, and W. H. Mangione-Smith, "The filter cache: an energy efficient memory structure," in *30th Annu. ACM/IEEE Int. Symp. on Microarchitecture*, 1997, Conference Proceedings, pp. 184–193.
- [49] N. Bellas, I. Hajj, C. Polychronopoulos, and G. Stamoulis, "Energy and performance improvements in microprocessor design using a loop cache," in *Computer Design, (ICCD '99), International Conference on*, 1999, Conference Proceedings, pp. 378–383.
- [50] L. Lea Hwang, B. Moyer, and J. Arends, "Instruction fetch energy reduction using loop caches for embedded applications with small tight loops," in *Low Power Electronics and Design*, 1999, Conference Proceedings, pp. 267–269.
- [51] A. Gordon-Ross, S. Cotterell, and F. Vahid, "Exploiting fixed programs in embedded systems: A loop cache example," *IEEE Computer Architecture Letters*, vol. 1, no. 1, p. 2, 2002.
- [52] B. Egger, J. Lee, and H. Shin, "Scratchpad memory management for portable systems with a memory management unit," in *Proceedings of the 6th ACM & IEEE International Conference on Embedded Software*, ser. EMSOFT '06. New York, NY, USA: ACM, 2006, pp. 321–330. [Online]. Available: <http://doi.acm.org/10.1145/1176887.1176933>

- [53] M. Verma, L. Wehmeyer, and P. Marwedel, "Cache-aware scratchpad allocation algorithm," in *Design, Automation and Test in Europe Conf. and Exhibition*. IEEE Comput. Soc, 2004, Conference Proceedings, pp. 1264–1269.
- [54] S. Borkar, "Design challenges of technology scaling," *Micro, IEEE*, vol. 19, no. 4, pp. 23–29, Jul.-Aug. 1999.
- [55] C. Guangyu, L. Feihui, O. Ozturk, C. Guilin, M. Kandemir, and I. Kolcu, "Leakage-aware SPM management," in *IEEE Computer Society Annu. Symp. on Emerging VLSI Technologies and Architectures*, 2006, Conference Proceedings.
- [56] M. Kandemir, M. J. Irwin, G. Chen, and I. Kolcu, "Compiler-guided leakage optimization for banked scratch-pad memories," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 10, pp. 1136–1146, October 2005.
- [57] Y. Huangfu and W. Zhang, "Compiler-based approach to reducing leakage energy of instruction scratch-pad memories," in *Computer Design (ICCD), 2013 IEEE 31st International Conference on*, Oct 2013, pp. 439–442.
- [58] H. Takase, H. Tomiyama, G. Zeng, and H. Takada, "Energy efficiency of scratch-pad memory in deep submicron domains: an empirical study," *IEICE Electronics Express*, vol. 5, no. 23, pp. 1010–1016, 2008.
- [59] S. Thoziyoor, N. Muralimanohar, J. Ahn, and N. Jouppi. (2008) CACTI 5.1. [Online]. Available: <http://www.hpl.hp.com/techreports/2008/HPL-2008-20.html>
- [60] (2005) International Technology Roadmap for Semiconductors. [Online]. Available: <http://www.itrs.net/reports.html>
- [61] S. Wilton and N. Jouppi, "CACTI: An enhanced cache access and cycle time model," *IEEE J. Solid-State Circuits*, vol. 31, no. 5, pp. 677–688, May 1996.
- [62] C. H. Jan, P. Bai, J. Choi, G. Curello, S. Jacobs, J. Jeong, K. Johnson, D. Jones, S. Klopceic, J. Lin, N. Lindert, A. Lio, S. Natarajan, J. Neiryneck, P. Packan, J. Park, I. Post, M. Patel, S. Ramey, P. Reese, L. Rockford, A. Roskowski, G. Sacks, B. Turkot, Y. Wang, L. Wei, J. Yip, I. Young, K. Zhang, Y. Zhang, M. Bohr, and B. Holt, "A 65nm ultra low power logic platform technology using uni-axially strained-silicon transistors," in *IEDM Tech. Dig.*, 2005, Conference Proceedings, pp. 60–63.
- [63] C. H. Jan, P. Bai, S. Biswas, M. Buehler, Z. P. Chen, G. Curello, S. Gannavaram, W. Hafez, J. He, J. Hicks, U. Jalan, N. Lazo, J. Lin, N. Lindert, C. Litteken, M. Jones, M. Kang, K. Komeyli, A. Mezhiba, S. Naskar, S. Olson, J. Park, R. Parker, L. Pei, I. Post, N. Pradhan, C. Prasad, M. Prince, J. Rizk, G. Sacks, H. Tashiro, D. Towner, C. Tsai, Y. Wang, L. Yang, J. Y. Yeh, J. Yip, and K. Mistry, "A 45nm low power system-on-chip technology with dual gate (logic and i/o) high-k/metal gate strained silicon transistors," in *IEDM Tech. Dig.*, 2008, Conference Proceedings.
- [64] C. H. Jan, M. Agostinelli, M. Buehler, Z. P. Chen, S. J. Choi, G. Curello, H. Deshpande, S. Gannavaram, W. Hafez, U. Jalan, M. Kang, P. Kolar, K. Komeyli, B. Landau, A. Lake, N. Lazo, S. H. Lee, T. Leo, J. Lin, N. Lindert, S. Ma, L. McGill, C. Meining, A. Paliwal, J. Park, K. Phoa, I. Post, N. Pradhan, M. Prince, A. Rahman, J. Rizk, L. Rockford, G. Sacks, A. Schmitz, H. Tashiro, C. Tsai, P. Vandervoorn, J. Xu, L. Yang, J. Y. Yeh, J. Yip, K. Zhang, Y. Zhang, and P. Bai, "A 32nm soc platform technology with 2nd generation high-k/metal gate transistors optimized for

- ultra low power, high performance, and high density product applications,” in *IEDM Tech. Dig.*, 2009, Conference Proceedings.
- [65] C. H. Jan, U. Bhattacharya, R. Brain, S.-J. Choi, G. Curello, G. Gupta, W. Hafez, M. Jang, M. Kang, K. Komeyli, T. Leo, N. Nidhi, L. Pan, J. Park, K. Phoa, A. Rahman, C. Staus, H. Tashiro, C. Tsai, P. Vandervoorn, L. Yang, J.-Y. Yeh, and P. Bai, “A 22nm SoC platform technology featuring 3-D tri-gate and high-k/metal gate, optimized for ultra low power, high performance and high density soc applications,” in *IEDM Tech. Dig.*, 2012, Conference Proceedings, pp. 3.1.1 – 3.1.4.
  - [66] R. Dennard, F. Gaensslen, V. Rideout, E. Bassous, and A. LeBlanc, “Design of ion-implanted MOSFET’s with very small physical dimensions,” *Solid-State Circuits, IEEE Journal of*, vol. 9, no. 5, pp. 256–268, Oct 1974.
  - [67] H. S. Yang, R. Wong, R. Hasumi, Y. Gao, N. S. Kim, D. H. Lee, S. Badrudduza, D. Nair, M. Ostermayr, H. Kang, H. Zhuang, J. Li, L. Kang, X. Chen, A. Thean, F. Arnaud, L. Zhuang, C. Schiller, D. P. Sun, Y. W. Teh, J. Wallner, Y. Takasu, K. Stein, S. Samavedam, D. Jaeger, C. V. Baiocco, M. Sherony, M. Khare, C. Lage, J. Pape, J. Sudijono, A. L. Steegen, and S. Stiffler, “Scaling of 32nm low power SRAM with high-k metal gate,” in *IEDM Tech. Dig.*, 2008, Conference Proceedings.
  - [68] K. Fujita, Y. Torii, M. Hori, J. Oh, L. Shifren, P. Ranade, M. Nakagawa, K. Okabe, T. Miyake, K. Ohkoshi, M. Kuramae, T. Mori, T. Tsuruta, S. Thompson, and T. Ema, “Advanced channel engineering achieving aggressive reduction of  $V_T$  variation for ultra-low-power applications,” in *IEDM Tech. Dig.*, 2011, Conference Proceedings, pp. 32.3.1–32.3.4.
  - [69] P. Athe and S. Dasgupta, “A comparative study of 6T, 8T and 9T decanano SRAM cell,” in *IEEE Symp. on Industrial Electronics & Applications (ISIEA)*, vol. 2, 2009, Conference Proceedings, pp. 889–894.
  - [70] B. H. Calhoun and A. Chandrakasan, “A 256kb sub-threshold SRAM in 65nm CMOS,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)*, 2006, Conference Proceedings, pp. 2592–2601.
  - [71] S. Lin, Y.-B. Kim, and F. Lombardi, “A low leakage 9t sram cell for ultra-low power operation,” in *Proc. of the 18th ACM Great Lakes Symp. on VLSI*. New York, New York, USA: ACM Press, 2008, Conference Proceedings, pp. 123–126.
  - [72] E. Marsman, R. Senger, and M. McCorquodale, “A 16-bit low-power microcontroller with monolithic MEMS-LC clocking,” in *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, May 2005, Conference Proceedings, pp. 624–627.
  - [73] N. Magen, A. Kolodny, U. Weiser, and N. Shamir, “Interconnect-power dissipation in a microprocessor,” in *Proceedings of the 2004 International Workshop on System Level Interconnect Prediction*, ser. SLIP ’04. New York, NY, USA: ACM, 2004, pp. 7–13. [Online]. Available: <http://doi.acm.org/10.1145/966747.966750>
  - [74] (2005) Predictive Technology Model. [Online]. Available: <http://ptm.asu.edu/interconnect.html>
  - [75] (2011) International Technology Roadmap for Semiconductors. [Online]. Available: <http://www.itrs.net/reports.html>

- [76] C. Lee, M. Potkonjak, and W. Mangione-Smith, "Mediabench: a tool for evaluating and synthesizing multimedia and communications systems," in *Thirtieth Annual IEEE/ACM International Symposium on Microarchitecture*, Dec 1997, pp. 330–335.
- [77] M. Guthaus, J. Ringenberg, D. Ernst, T. Austin, T. Mudge, and R. Brown, "Mibench: A free, commercially representative embedded benchmark suite," in *Workload Characterization, 2001. WWC-4. 2001 IEEE International Workshop on*, Dec 2001, pp. 3–14.
- [78] A. Steegen, R. Mo, R. Mann, M.-C. Sun, M. Eller, G. Leake, D. Vietzke, A. Tilke, F. Guarin, A. Fischer, T. Pompl, G. Massey, A. Vayshenker, W. L. Tan, A. Ebert, W. Lin, W. Gao, J. Lian, J.-P. Kim, P. Wrschka, J.-H. Yang, A. Ajmera, R. Knoefler, Y.-W. Teh, F. Jamin, J. Park, K. Hooper, C. Griffin, P. Nguyen, V. Klee, V. Ku, C. Baiocco, G. Johnson, L. Tai, J. Benedict, S. Scheer, H. Zhuang, V. Ramanchandran, G. Matusiewicz, Y.-H. Lin, Y. Siew, F. Zhang, L. S. Leong, S. L. Liew, K. Park, K.-W. Lee, D. H. Hong, S.-M. Choi, E. Kaltalioglu, S. Kim, M. Naujok, M. Sherony, A. Cowley, A. Thomas, J. Sudijohno, T. Schiml, J.-H. Ku, and I. Yang, "65nm CMOS technology for low power applications," in *Electron Devices Meeting, IEDM Technical Digest. IEEE International*, Dec 2005, pp. 64–67.
- [79] Lakkasuo, "Diagram of a push-pull amplifier," Sept 2010. [Online]. Available: [http://en.wikipedia.org/wiki/File:Electronic\\_Amplifier\\_Push-pull.svg](http://en.wikipedia.org/wiki/File:Electronic_Amplifier_Push-pull.svg)
- [80] R. Hogervorst, *Design of low-voltage, low-power operational amplifier cells*. Springer, 1996.
- [81] K. Kraver, "Low-voltage and low-power, deep submicron analog circuits for a single-chip mixed-signal microinstrumentation system," Ph.D. dissertation, University of Michigan, 2003.
- [82] B. Zhai, S. Hanson, D. Blaauw, and D. Sylvester, "Analysis and mitigation of variability in subthreshold design," in *Low Power Electronics and Design, 2005. ISLPED '05. Proceedings of the 2005 International Symposium on*, Aug 2005, pp. 20–25.
- [83] M. Steyaert, V. Peluso, J. Bastos, P. Kinget, and W. Sansen, "Custom analog low power design: the problem of low voltage and mismatch," in *Custom Integrated Circuits Conference, 1997., Proceedings of the IEEE 1997*, May 1997, pp. 285–292.
- [84] B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw-Hill, 2001.
- [85] D. A. Johns and K. Martin, *Analog Integrated Circuit Design*. John Wiley & Sons, 1997.
- [86] D. O. Wipf, F. Ge, T. W. Spaine, and J. E. Baur, "Microscopic measurement of pH with iridium oxide microelectrodes," *Analytical Chemistry*, vol. 72, no. 20, pp. 4921–4927, 2000. [Online]. Available: <http://dx.doi.org/10.1021/ac000383j>
- [87] L. Bin, T. W. Rondeau, J. H. Reed, and C. W. Bostian, "Analog-to-digital converters," *Signal Processing Magazine, IEEE*, vol. 22, no. 6, pp. 69–77, 2005.
- [88] L. Hae-Seung and C. G. Sodini, "Analog-to-digital converters: Digitizing the analog world," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 323–334, 2008.

- [89] K. Abdelhalim, L. MacEachern, and S. Mahmoud, "A nanowatt ADC for ultra low power applications," in *Circuits and Systems, IEEE International Symposium on (ISCAS)*, 2006, Conference Proceedings, p. 4.
- [90] W. Chun-Hao and L. Liang-Hung, "A 0.6-V delta-sigma ADC with 57-dB dynamic range," in *VLSI Design Automation and Test (VLSI-DAT), International Symposium on*, 2010, Conference Proceedings, pp. 198–201.
- [91] A. Matsuzawa, "High speed and low power ADC design with dynamic analog circuits," in *IEEE 8th International Conference on ASIC (ASICON)*, 2009, Conference Proceedings, pp. 218–221.
- [92] J. K. Fiorenza, T. Sepke, P. Holloway, C. G. Sodini, and L. Hae-Seung, "Comparator-based switched-capacitor circuits for scaled CMOS technologies," *Solid-State Circuits, IEEE Journal of*, vol. 41, no. 12, pp. 2658–2668, 2006.
- [93] L. Brooks and L. Hae-Seung, "A zero-crossing-based 8-bit 200 MS/s pipelined ADC," *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 12, pp. 2677–2687, 2007.
- [94] C. Youngcheol and H. Gunhee, "Low voltage, low power, inverter-based switched-capacitor delta-sigma modulator," *Solid-State Circuits, IEEE Journal of*, vol. 44, no. 2, pp. 458–472, 2009.
- [95] Q. Da, L. Yuan-wen, C. Long, X. Jun, Y. Fan, and R. Jun-Yan, "An ultra low power sigma-delta modulator for hearing aid with double-sampling," in *IEEE 8th International Conference on ASIC (ASICON)*, 2009, Conference Proceedings, pp. 1141–1144.
- [96] V. S. L. Cheung, H. Luong, and M. Chan, "A 0.9-V 0.2  $\mu$ W CMOS single-opamp-based switched-opamp  $\Sigma\Delta$  modulator for pacemaker applications," in *Circuits and Systems, IEEE International Symposium on (ISCAS)*, vol. 5, 2002, Conference Proceedings, pp. 185–188.
- [97] A. Fazli Yeknami, F. Qazi, J. J. Dabrowski, and A. Alvandpour, "Design of OTAs for ultra-low-power sigma-delta ADCs in medical applications," in *Signals and Electronic Systems (ICSES), International Conference on*, 2010, Conference Proceedings, pp. 229–232.
- [98] R. Schreier and G. Temes, *Understanding Delta-Sigma Data Converters*. Wiley, 2004. [Online]. Available: <http://books.google.com/books?id=5HNqQgAACAAJ>
- [99] B. Boser and B. Wooley, "The design of sigma-delta modulation analog-to-digital converters," *Solid-State Circuits, IEEE Journal of*, vol. 23, no. 6, pp. 1298–1308, Dec 1988.
- [100] E. Hogenauer, "An economical class of digital filters for decimation and interpolation," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, no. 2, pp. 155–162, 1981.
- [101] S. H. Tan, "Simulating CIC filtering/decimation without the filter design toolbox," 2009. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/25590-simulating-cic-filteringdecimation-without-the-filter-design-toolbox/>
- [102] R. Schreier, "Delta Sigma Toolbox," 2000. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/19-delta-sigma-toolbox>

- [103] T. Saramaki, "Design of FIR filters as a tapped cascaded interconnection of identical subfilters," *Circuits and Systems, IEEE Transactions on*, vol. 34, no. 9, pp. 1011–1029, 1987.
- [104] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*. Oxford University Press, 2011.
- [105] R. J. Van de Plassche, *CMOS Integrated Analog-to-Digital and Digital-to-Analog Converters*. Kluwer Academic Publishers Dordrecht, 2003.
- [106] A. P. Chandrakasan and R. W. Brodersen, *Low Power CMOS Digital Design*. Springer Science+Business Media, LLC, 1995.
- [107] H. Schouwenaars, D. Groeneveld, and H. A. H. Termeer, "A low-power stereo 16-bit CMOS D/A converter for digital audio," *Solid-State Circuits, IEEE Journal of*, vol. 23, no. 6, pp. 1290–1297, Dec 1988.
- [108] D. Mercer, "Low-power approaches to high-speed current-steering digital-to-analog converters in 0.18  $\mu\text{m}$  CMOS," *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 8, pp. 1688–1698, Aug 2007.
- [109] J. Jamp, J. Deng, and L. Larson, "A 10GS/s 5-bit ultra-low power DAC for spectral encoded ultra-wideband transmitters," in *Radio Frequency Integrated Circuits (RFIC) Symposium, 2007 IEEE*, June 2007, pp. 31–34.
- [110] S. Roundy, P. K. Wright, and J. Rabaey, "A study of low level vibrations as a power source for wireless sensor nodes," *Computer Communications*, vol. 26, no. 11, pp. 1131–1144, 2003.